Research Intelligence

B.E.R.A. Bulletin 2

Stirling 75

CONTENTS

STERLING INTELLIGENCE

This issue is devoted to the proceedings of the 1975 BERA conference at Stirling. The sessions comprised a wide range of contributions from several disciplines within Education and included papers presented to plenary sessions and introductory papers to symposia. Most contributions are printed here and comprise papers concerned with experimental matters, with conceptual and philosophical issues, with historical perspectives and with the business of research. There are various styles and emphases but we have chosen not to impose a uniform format on them.

George Brown Neville Bennett
Mike Smith Ed Stones

Printed & designed by Birmingham Arts Lab, 021-359 4192.

Cover design by Ed Stones.

ADDRESSES TO CONFERENCE

THE COLOUR OF CONCEPTUAL LEARNING

Edgar Stones.
Liverpool University.

INTRODUCTION

In this paper I should like to discuss some aspects of psychology and education that have interested me for a number of years but which have recently attracted particular attention through the work of Jensen. I refer to questions concerned with children's learning and the problem of improving this learning in school. This pedagogical orientation is, of course, different from Jensen's main preoccupation which seems to be with ethnic differences in learning ability or intelligence. However, he has suggested pedagogical procedures on the basis of his work and it is this aspect of his work that I wish to focus on.

Jensen has argued that there are two different types of learning ability which he calls level I and level II. These correspond roughly to rote and conceptual learning. While level I is distributed similarly in different populations, level II is distributed differently. Children with white faces seem, according to Jensen, to have the monopoly of level II ability. Jensen argues that these abilities are distributed in this way according to genetic laws and that inherited factors are the most important in determining an individual's learning ability. Various critics have taken up Jensen's arguments with varying degrees of acerbity and have dealt with such things as the theoretical foundations of his arguments about heritability, the validity of the statistical underpinning to his thesis and his conceptualising of the nature of intelligence. One commentator has also drawn attention to Jensen's frequent misquoting of evidence to back up his thesis. (Deutsch 1973).

It is not my wish, to join in the debate about nature versus nurture or about the statistical validity of Jensen's argument, but I would like to consider what seems to me to be quite fundamental to the whole issue, that is the nature of the abilities under question and the means adopted to assess them. I am not referring to the question of intelligence testing, but to children's learning abilities

During the past ten to fifteen years the study of school learning has increasingly attracted the attention of academic psychologists and there has been a corresponding de-emphasis on studies in the field of intelligence testing. It is, therefore, somewhat surprising that it was necessary for Deutsch to remind us that the view of intelligence adopted by Jensen (the one proposed by Spearman in 1923) is only one among many theories none of which have been 'proven'; not to mention the fact that intelligence tests measure essentially what children have learned, not how well they might learn something new. It is the latter point that I wish to address myself to because it seems to me that Jensen misconceives the nature of the learning he claims to be investigating and the nature of instruments that might assess that learning. I think that the instruments he uses do not operate in the way he argues and wish to examine his rationale for using them and to adduce some experimental evidence that bears on the issue.

ASSESSING ABILITY

Studies of cognitive processes in different ethnic groups have, in the main, focussed on performance on standardised tests or specially constructed test material designed to test existing competence in a variety of fields. Various attempts have been made to devise 'culture fair' tests to allow of meaningful comparisons among different cultural groups. (Anastasi 1965, Lesser et al 1967). There has been some disagreement about the validity of the concept of 'culture fairness' and it has been asserted that it is virtually impossible to devise a truly culture fair test. (Wesman 1968). However, Jensen (1973b) has argued at length that it is possible to look upon some nonverbal tests of intelligence as being what he refers to as 'status free'. From among these tests he singles out the Raven's matrices test as the most appropriate measure.

Among the reasons he gives for supporting the Raven's test is that there have been no studies that demonstrate gains in relatively noncultural or nonverbal tests like Cattell's Culture Fair Tests or Raven's Matrices. This is not strictly true, although I suppose we can forgive Jensen for not being acquainted with the findings of an unpublished master's thesis in an obscure European university. In fact a study carried out by Renhard (Renhard 1971) under my direction not only addresses itself to this question but also exemplifies the general approach adopted in the empirical studies reported in this paper and, I believe, develops a more appropriate approach to school learning than does that of Jensen.

Our point of departure in this study was to examine the extent to which performance in the Raven's test could be improved by teaching. It is important to stress that we were interested in teaching not coaching on the test itself. We adopted a self instructional programmed learning approach and set as our criterion of success the ability to achieve a higher score on the matrices after teaching than at the beginning. However, the instruction given to the children at no time made use of the actual Raven's material. Instead a teaching programme was devised that made use of a variety of materials as specific exemplars of the principles behind the matrices. Some of these materials were pictorial, some were abstract and some were in the form of numbers. At no time was there any attempt to drill. The examples were programmed with the intention of providing a sufficient variety of exemplars and non exemplars for the children to acquire the necessary concepts. In other words, the teaching was much in line with the approach that one would adopt in helping children to learn new concepts or principles in school teaching. The criterion test was the Raven's matrices test. After using the self instructional programme for approximately $2\frac{1}{2}$ hours children in the first year of secondary education in an industrial town made significant gains on the matrices test as compared with a control group that did not work through the programme.

Apart from the evidence of gains following the use of the programme, there is an important point in the way the Raven's test was used. In this experiment it was used as a test of learning related to specific teaching or learning activities. In other words we were not attempting to sample a hypothesised ability but to see to what extent the mental skills involved in coping with the problem were influencable by teaching.

Jensen, (1968, p.1331) himself, would seem to be in favour of this approach since he argues that "since standard intelligence tests contain items intended to assess how much the individual has learned in his natural environment, a more direct and relatively culture-free index of intelligence might be the rate or amount of learning in a novel laboratory task". He has, in a variety of studies, adopted this approach to the assessment of what he terms level I abilities, i.e. what might be generally referred to as rote learning tasks, and he has presented children with such tasks as paired associate learning and digit span. However, when he has turned to his so called 'level II abilities he has not used a comparable complex-learning task to assess the ability to cope with conceptual learning. Instead he has made considerable use of the Raven's matrices test, not in the way that Renhard used it, as a test of learning related to a specific learning task, but as a general indicator of the ability to cope with higher mental processes that depend "... upon elaborations and transformations of informational input, and upon comparisons of the informational input with previously stored information". After extensive test programmes he found that the children of low SES status from different ethnic backgrounds obtained similar scores to high SES white children on rote learning tasks such as paired associate learning and digit span. However, low SES children made lower scores on standard intelligence tests. He found low correlations between intelligence scores derived from the Raven's Progressive Matrices, and rote learning scores in low SES populations but substantial correlations among the measures in high SES pop-populations. In attempting to account for these differences he advanced three hypotheses.

a) There are two genotypically independent cognitive processes: one (level I) appropriate to rote learning and one (level II) appropriate to the solution of Matrix problems which involve abstraction, generalisation and symbolisation.

b) Level II processes are functionally dependent upon level I. That is the growth rate and asymptote of a child's performance on level II depends on his status on level I.

c) Level I ability is distributed approximately the same in all socio economic classes. Level II ability is distributed about a higher mean in the higher socio economic classes.

Although Jensen points out that in the experiments which led to these hypotheses, race was confounded with SES, he argues later that 'Because short term achievement measures reflect factors other than intelligence, Negroes and whites differ slightly less on such measures than they differ on intelligence tests. ... The problem of Negro-white inequality in educability is thus essentially the problem of Negro-white differences in intelligence'. (Jensen 1973a p. 355).

I think that Jensen's confounding of SES and ethnic group is very much open to question. Ghuman, in a very thorough investigation of the ability of children in different ethnic groups found that the cultural and socioeconomic background of the children was the crucial factor in achievement at different levels of learning. Children of the same ethnic group in their native country performed differently from their counterparts in England and those in England performed similarly to their English peers. (Ghuman 1974).

However, Jensen does not hesitate to argue that his analyses have important educational implications. Traditionally (he avers) schools have been organised to employ teaching methods based on conceptual learning (his type II ability) and therefore have disadvantaged low SES and particularly negro children who learn better by type I associative methods. Schools should therefore, address themselves to developing methods of teaching which capitalise on the rote learning abilities of these children and '...provide thereby a means of improving the educational attainments of many of the children now called culturally disadvantaged'. (Jensen 1968 p.1337).

I believe there are several issues here that are open to question. I wonder, for example, how his assertion that schools employ teaching methods based on conceptual learning, would stand up to scrutiny. No doubt some do but I suggest it is a bold stance to assert that this is the method, as he seems to do. Although I believe this question is very important and one which would well repay investigation, it is not the main focus of my discussion. My concern is to question the basis upon which his recommendations are founded.

The aspect of Jensen's work that I want to examine is his use of rote learning tasks to assess his hypothesised type I ability and conceptual tests to assess his type II ability. As I suggested earlier when referring to Renhard's work, Jensen measures rote learning direct but conceptual learning by inference. The former samples competence in new learning, i.e. learning new stimulus combinations, the latter samples existing competence related to the solution of the matrices problems when the only learning allowed for is in the trial items I would argue that this is a weakness in his approach that needs careful scrutiny and believe that a more appropriate method of investigating ability in any supposed different types of learning, would be to set up different types of learning tasks and compare performance on these. Jensen's approach to assessing rote learning seems satisfactory but it seemed to me essential to employ some kind of concept learning task followed by a concept learning test in order to sample performance in a conceptual learning task rather than existing competence in an intelligence test.

A very similar point made by Deutsch (1973) bears on the question of the approach to the investigation. This is that Jensen's approach has been entirely psychometric: to the extent that the problem is worthy of investigation other approaches should be adopted. I should like now to describe some work I have carried out with the help of teacher colleagues, that attempted to provide two kinds of learning task in experimental situations and which I hoped would make possible a more meaningful appraisal of the supposed differences in learning abilities with particular reference to different ethnic groups.

THE EXPERIMENTAL STUDIES

The aim was to investigate experimentally the actual learning of different groups of children in tasks involving rote learning and conceptual learning. The test of rote learning was of the same genre as paired associates and was thus comparable with some of Jensen's instruments. On the other hand the conceptual learning task was a genuine test of learning and was quite different from Jensen's use of the Raven's matrices. The learning task used the Vigotsky concept learning apparatus and transfer test equipment that I had designed specially for use in earlier investigations into the processes of learning fairly complex concepts by primary school children. (Stones and Heslop 1968, Stones 1970).

The conceptual learning apparatus was originally used by Vigotsky as a method of investigating the role of language in concept learning. (Vigotsky 1962) He was interested in the actual processes of concept learning. Subsequent users of the apparatus have employed it as a test of intelligence. (Semeonoff and Laird 1952, Hanfmann & Kassinin 1937). None of these made use of transfer tests nor is there any record of such tests being used by other workers. Vigotsky examined the way subjects manipulated the experimental material in an attempt to investigate the underlying cognitive processes, and testers have used the material as a classificatory sorting test of existing cognitive competence using as indices of success such things as number of unsuccessful groupings before solution and time taken to complete the sorting. In my use of the apparatus I have used the Vigotsky sorting task as a learning task and the transfer test equipment as material for investigating whether or not the children had actually acquired the concepts underlying ability to succeed in the Vigotsky task. Thus the novel feature of these studies and those reported here is that the grouping of the blocks is not scored but treated as a learning experience in the course of which the subject learns new concepts. The test of the learning is success in classifying novel material according to the same criteria. In assessing rote learning, this form of transfer test is not appropriate since the arbitrary nature of the connexions made precludes transfer, whereas in conceptual learning the acid test of success is the ability to transfer.

THE APPROACH

In order to examine children's learning two learning tasks were devised. One of these was a rote learning task with a specified level of competence as the criterion of success. The other was the conceptual learning task with a transfer test of competence. The Raven's nonverbal test of intelligence was also used.

Apparatus

The apparatus used for the rote learning task consisted of five cards each with a simple geometrical shape on the one side and a digit on the reverse. The shapes comprised circle, rectangle, triangle, semi-circle, and elongated rectangle with one rounded end. The digits on the reverse were arbitrarily determined as 8, 6, 3, 4 and 7.

The apparatus used for the concept learning task was the standard Vigotsky blocks. These comprise twenty-two small wooden blocks of five different colours, six different cross sections, two different heights or thicknesses, and two different cross sectional areas. Under each block is written one of four nonsense words: LAG, BIK, MUR, and CEV. The problem is to classify the blocks in four groups so that all the blocks in any one group have common properties which unequivocally mark them off as members of that group and as nonmembers of any other group. The criterial attributes for 'correct' grouping are cross-sectional area and height. The nonsense syllables relate to the criterial attributes, and may, in fact, be considered as 'names' of the concepts exemplified by the blocks. Thus, the LAG blocks are tall and fat, the BIK blocks are small and fat, the CEV blocks are small and thin, and the MUR blocks are tall and thin. There is no other consistent way of making four groups of the blocks.

In the experiment the blocks are spread at random on a table, nonsense words down. The learning task is to arrange the blocks in the four groups. The child is shown an example by the experimenter's turning over one of the blocks and showing the nonsense word. The child is asked to arrange the blocks in the four groups and is told that he is free to move blocks from one group to another or back to the pool of unsorted blocks whenever he wishes. He is not allowed to turn over any of the blocks. Whenever the child completes a sorting incorrectly, the experimenter turns over one incorrect block grouped with the original specimen and shows that it is different from the specimen and encourages the child to try

again. With each wrong sorting the number of upturned blocks increases and provides additional clues to the child. Thus as the number of upturned blocks increases the child is able to obtain a basis for discovering to which characteristics of the blocks the nonsense words refer. When the child makes this discovery he is able to complete the task with facility.

The apparatus used to test the children's learning of the concepts involved in successful sorting of the Vigotsky material, comprised seventeen small objects which could be classified according to the same criteria as the Vigotsky blocks. The objects were sufficiently dissimilar in size, shape and colour from the blocks to preclude transfer by primary stimulus generalisation. They comprised the following objects which were of a variety of colours, shapes and materials.

Group 1 LAG (Tall and fat): cardboard box, irregular plaster block, toy building brick, plastic mug.

Group 2 MUR (Tall and thin): candle, cardboard cylinder, torch battery, toy giraffe.

Group 3 (BIK (small and fat): rubber wash plug, flat square plastic block, flat plastic disc, tablet of soap, matchbox.

Group 4 CEV (small and thin): packet of chewing gum, pencil sharpener, die ring.

The validity of the grouping was checked by trying out the apparatus with adults who knew the Vigotsky material and no disagreement was found with respect to the sorting of the objects in the different categories.

In addition to the two tests of learning, the Raven's test of nonverbal intelligence was used. Two versions were employed, the Coloured Progressive Matrices (Raven 1969) for the younger children, and the Standard Progressive Matrices (Raven 196O) for the older children. The Matrices are held to '...provide five opportunities for grasping the method and five progressive assessments of a person's capacity for intellectual activity'. (Raven 196O). It should be noted that this test, unlike the Vigotsky learning task, does not allow for the subject to receive any feed-back as to the correctness of his efforts.

THE CHILDREN

Study 1

In this first investigation 3O children of West Indian origin and 3O white children (15 boys and 15 girls in both cases) were given the learning tasks and the Raven test. The sample was drawn with the classroom as the unit of selection in such a way as to include all the children of the ethnic group with fewer members in the class and an equivalent number of the other ethnic group drawn randomly. The children came from two inner ring primary schools in a large industrial city and were aged between ten and eleven years. Almost all the children lived in municipal housing in the vicinity of the school so that although no attempt was made to measure socio economic status, it seems reasonable to assume that the sample was homogeneous with regard to SES.

Study 2

This investigation was a replication of the first study and adopted exactly the same procedure except that on this occasion the children were aged 13 to 14 and were in second year of a secondary school. Twenty three white children on this occasion were compared with 23 Pakistani children in the same classes and from the same neighbourhood. There were 15 white and 15 Pakistani boys and 8 white and 8 Pakistani girls.

THE EXPERIMENTAL TASKS (BOTH STUDIES)

The rote learning task

This task is considered to be rote learning because there is no logical or conceptual relationship between the two stimuli, the geometrical shape and the digit. The connexion between the two is arbitrary as in paired associated learning experiments. The child has to learn the connexion between the shapes and the digits.

The tasks were administered individually in a quiet room. After rapport had been established, the experimenter presented the experiment in the form of a game. He explained to the child that each card has a shape on one side and a number on the other. The child has to learn the number that goes with each shape. The experimenter shows the first card, shape towards the child, and then turns it over to expose the digit and says the number. The experimenter repeats the process with the remaining four cards. He then shuffles the cards and exposes the first card, shape towards the child, and asks him to give its number. Whatever answer the child gives, the experimenter turns the card to expose the digit.

The experimenter repeats the process with the remaining four cards. He then repeats the procedure until the child gives the correct numbers when shown the shapes for three successive trials of five cards each. The experimenter shuffles the cards between each trial. The index of learning is number of trials taken to criterion: the lower the number of trials the quicker the learning.

The concept learning task

This task was conducted in the same session as the rote learning task and was presented as a game in which the child had to find out the basis of grouping the blocks that the experimenter has decided upon. No suggestion was made that the groupings decided upon were the 'correct'groupings. Apart from encouraging the child to keep on trying no attempt was made to teach or help in any way. The learning task was to make use of feedback from the words on upturned wrong blocks to discover the basis of grouping which was according to the standard criteria of height and size of cross section. (See Stones and Heslop 1968).

The first step was for the child to group the blocks correctly. When a child had grouped the blocks correctly they were brushed up and presented again to the child for regrouping. On this occasion only one attempt was allowed. The initial operations in discovering the basis of grouping involved, in most cases, several incorrect attempts followed by feedback from the wrongly placed upturned blocks. This phase of the child's activity was considered not as a test but as a learning activity. The first correct grouping was taken to be a possible indication that the child had learned the basis of grouping. The regrouping was taken to be a test of the retention of the learning. Finally the 17 heterogenous objects were presented as a transfer test of the learn ing after the blocks had been removed from sight. The child is asked to put the objects in the same groupings as the blocks. Successful grouping of these objects was taken to be the most reliable evidence of the learning of the conceptual basis of the grouping operations. The data collected in this task were success or failure in the initial grouping, success or failure in regrouping the blocks and success in sorting the new objects correctly. The criterion of success in each case was completely accurate grouping.

The nonverbal tests of intelligence.

The Raven's matrices were administered as group tests to all the subjects. Several experienced teachers familiar with the administration of the test were present to ensure that the children were clear about the nature of the task.

RESULTS

The rote learning task.

All groups of children in the two studies performed similarly in this test. The number of trials to criterion ranged from three which was immediate learning after the demonstration, to a discontinued test after 28 trials. The average number of trials for all groups was about 11. To investigate possible differences in performance by the different groups of children, the scores were split into high and low at the median and chi square analysis carried out. There were no significant differences in performance between the two groups. Full details are given in tables 1 and 2.

TABLE 1 ROTE LEARNING TASK: STUDY 1

Children	Trials to criterion	
	Mean	S.D.
West Indian	12.31	5.85
White	10.91	5.14
Total Group	11.61	5.66

	High/Low Split	
	High	Low
West Indian	18	12
White	15	15

Chi square = 0.61. (N.S.)

TABLE 2 ROTE LEARNING TASK: STUDY 2

Children	Trials to criterion	
	Mean	S.D.
Pakistani	11.80	4.1
White	11.66	5.12
Total Group	11.73	4.73

	High/Low Split	
	High	Low
Pakistani	13	10
White	10	13

Chi square = 0.78 (N.S.)

The concept learning task

A difference was observed in the performance of the children in the two studies. Children in the second study, who were about two years older than the ones in the first study, were more successful overall. Whereas 7 (out of 60) in study one failed to group the blocks successfully in the learning task, all 46 succeeded in the second study. Similarly a larger proportion of the total group in the second study succeeded in regrouping correctly and transferring their learning to the test material. In study one, the only marked difference in performance on the transfer test was between West Indian boys and White girls. This difference shows up in the black/white differences in transfer with ten West Indians succeeding as opposed to 16 whites, in both cases out of a possible 30. On the other hand, the West Indian children were more successful than the whites at regrouping. In study two there is hardly any discernible difference between the performance of the Pakistani and white children. To investigate further the differences between the performance of West Indian and White children chi square analysis was carried out dividing the groups into those who succeeded in transferring their learned ability and those who did not. The performance of Pakistani and White were so similar that significance tests were unnecessary. Full details may be seen in tables 3 and 4.

TABLE 3: PERFORMANCE OF CHILDREN ON THE CONCEPT LEARNING TASK: STUDY 1

	Failed	Grouped correctly only	Grouped and regrouped correctly	Grouped & regrouped correctly and transferred correctly
West Indian	4	2	14	10
White	3	7	4	16

Difference between the groups on successful transfer is not significant (Chi-square = 2.44).

TABLE 4: PERFORMANCE OF CHILDREN ON THE CONCEPT LEARNING TASK: STUDY 2

	Failed	Grouped correctly only	Grouped and regrouped correctly	Grouped & regrouped correctly and transferred correctly
Pakistani	0	1	6	16
	0	0	8	15

Significance tests on these data were considered unnecessary.

The nonverbal intelligence test (Raven's matrices)

Mean scores for children in study one were well below the average according to the published norms for the test and were all roughly about the 25th percentile. The children in study 2 were somewhat below the published norms but not to the same extent as in study one. In study one the West Indian children did slightly better than the white children and the white children did slightly better than the Pakistani children in study 2. There were, however, no significant differences among any of the groups in this test. Full details of mean scores may be seen in tables 5 and 6. As in the case of the rote learning scores a high/low split was carried out and the distributions checked for significance by calculation of chi square.

TABLE 5: SCORES ON THE RAVEN'S MATRICES TEST:

	Mean	S.D.
West Indian	24.14	5.48
White	23.09	6,80

	High/Low Split	
	High	Low
West Indian	12	18
White	15	15

Chi square = .61 (N.S.)

TABLE 6: SCORES ON THE RAVEN'S MATRICES TEST: STUDY 2

	Mean	S.D.
Pakistani	41.74	10.80
White	43.57	6.94

	High/Low Split	
	High	Low
Pakistani	10	13
White	10	13

Chi square = 0

Correlations among task performance and matrices scores.

In addition to the analyses of performance in the two learning tasks and the matrices test, contingency coefficients (C) were calculated among the three sets of scores for the combined groups (106 children). Raven's score was correlated with number of trials to criterion in the rote test and with the concept learning test but the concept learning test scores did not correlate with the rote scores. None of the correlations reached significant levels.

TABLE 7: CORRELATION (CONTINGENCY COEFFICIENT) AMONG TASK PERFORMANCE AND MATRICES SCORES: ALL CHILDREN (N = 106)

		1	2	3
1	Concept learning (Vigotsky)		0	.141
2	Rote learning			.149
3	Raven's score			

From among the results presented above, data relating to the overall comparisons between blacks and whites have been set out in figure 1 to present a synoptic overview of the relative performances of the different ethnic groups.

FIGURE 1: OVERALL COMPARISONS OF DIFFERENT GROUPS ON THE LEARNING TASKS AND THE NONVERBAL INTELLIGENCE TEST

A: Rote learning task; number of errors to criterion (the shorter the line the quicker the learning

White (1)

West Indian

White (2)

Pakistani

0 20

B: Conceptual learning task: percentage of children succeeding on transfer test

White (1)

West Indian

White (2)

Pakistani

0 100

C: Raven's scores (raw marks)

White (1)

West Indian

White (2)

Pakistani

0 30 (Coloured matrices)

60 (Standard matrices)

DISCUSSION

Clearly the most interesting thing that emerges from this investigation is the fact that children of different ethnic groups with similar socio-economic backgrounds in English schools do remarkably similarly on tests of learning, both rote and conceptual. The results resemble Jensen's insofar as rote learning is concerned. On the other hand, the instrument designed to assess conceptual learning as it happened, yielded very different results from those of Jensen. The low correlations between the test of conceptual learning reported here and the scores on the Raven's matrices are also of interest. If we accept Jensen's view that the Raven's test is a test of conceptual learning ability, and if you accept the test I have described also as a test of conceptual learning, then the low correlation between the two suggests that the two instruments may be identifying two different types of learning. We now have three types (levels?) of learning. The point, of course, is not to argue the relative merits of the Raven's test and the test used in my experiment, but that to suggest that human learning can be parcelled out into neat disjunctive categories is simplistic in the extreme.

The results obtained on the Raven's matrices test raise quite different issues. The two separate investigations reported here both fail to substantiate Jensen's findings of ethnic differences in scores on the Raven's test. However, it should be pointed out that Jensen's arguments relate essentially to American negro children and he suggests that other populations such as American Indians and Mexicans perform better than negroes but not so well as whites (Jensen, 1973, p.360). In his earlier writings on the subject of level I and level II abilities he laid more emphasis on the influence of SES. Later he argued that racial differences and differences in SES are so much related that matching groups for SES also matches them for genetic factors as well as he claims that the average skin colour of negroes becomes lighter in higher SES categories. (Jensen 1973, p.359). It is unlikely that these arguments have much relevance to the studies reported here. They could possibly have some bearing on the study of West Indian children but not at all on the results obtained from the study of Pakistani children. I suggest, therefore, that our investigations give a picture of ethnic comparisons in the learning ability of children in British schools and that those reported by Jensen are of little relevance to our conditions. The results of our studies, however, accord with those of Ghuman who found that different ethnic groups perform similarly if they have similar cultural backgrounds.

There is a further point about the conceptual learning task used in these studies. Apart from the fact that it attempts to assess learning that has actually taken place during rather than before the interview, it is a test of ability to cope with a specific learning task that can be objectively observed. It is not a norm referenced test devised on a statistical basis, it is a criterion referenced test, the criterion of success being the ability to apply a newly learned concept in a novel situation and not to obtain a higher score than the norm for one's age. There is more than academic interest in this difference between the tests. The pedagogical implications of the different approaches are profound. At best the Raven's score may be 'noted', or, if we follow Jensen's advice, we devise teaching procedures appropriate to the child's Raven's score; rote methods for the low scorers, 'conceptual methods' for the high scorers. I think we have been here before in this country in the immediate post war tripartism period with hewers of wood and drawers of water in different schools from the children capable of abstract thinking. If, however, we adopt the approach used in the Renhard study and implicit in the conceptual learning task I described, we draw very different pedagogical inferences. The test results in such cases are taken as diagnostic and indicators of necessary remedial action with the child for for the improvement of the teaching. Deutsch makes an important point when discussing the question of diagnosis in Jensen's arguments. Referring to the fact that Jensen enters frequent caveats with respect to not assuming a certain level on the part of any given individual on the basis of known group differences, he '...does not include any suggestions as to how one can identify a potential conceptual thinker in early childhood other than by his skin colour'. (Deutsch 1973 p.26).

CONCLUSIONS

In this paper I have looked at methods of investigating children's learning. In particular, I have scrutinised the methodology employed by Jensen in assessing the learning of children in different ethnic groups. I suggest that for this Association there are some important lessons to be learned from the way in which the original promulgation of the hypothesised two levels of learning were taken up and the line of argument in the subsequent controversy. Probably the most important is for us to consider the way in which research investigations and theoretical disputations of considerable sophistication can be built on simplistic views of pedagogy. The recherche debates about the exact proportion of ability that might be 'heritable' took as read Jensen's basic (and as I have contended, mistaken) views concerning the assessment of conceptual learning. Once those views are challenged the rest is noise.

Another example of Jensen's pedagogic naivety is when he assumes that schools teach for conceptual learning. Some teachers undoubtedly do, but their teaching is likely to be on an intuitive basis rather than on firmly grounded theoretical principles acquired in teacher training. It is true that experimental psychologists have acquuired some understanding of the way concepts are learned but we still have far to go. Further, there is no shortage of research to suggest that teachers in educational institutions at all levels are pristinely innocent of awareness of such evidence as we may have. And it is equally true that few training institutions have developed methods of teaching teachers to teach for conceptual learning in any explicit and systematic way.

In the work that I have reported I attempted to avoid what I saw as Jensen's errors and mistaken assumptions. Groups of children from complete classes were given the opportunity of learning fairly complex concepts from scratch and in these conditions different ethnic groups performed similarly. Thus the results of these studies lend no support to the thesis that there are racial differences in conceptual learning and I believe that the findings of these investigations would be applicable in comparable conditions, i.e. in typical British mixed school populations.

I suggest that the work I have reported challenges the validity of Jensen's diagnosis. I think we should all challenge his prescription however we view the evidence, because along with his presumption in claiming to be able to divide up humanity into two categories is an implicit assumption that we know all there is to know about teaching for conceptual learning. If this were not the case he could hardly argue with any consistency for concentrating on rote methods of teaching for his type I people unless he has other reasons for denying them the benefits of new pedagogical insights we might acquire. I think that few would argue that we have nothing more to learn about teaching for conceptual learning. I think this is far from the case and I suggest that we have so much to learn about teaching concepts to children of all ages that instead of trying to match the type of learning to the colour of the children we should all be far more fruitfully employed developing methods of enhancing the ability of children to cope with learning of all colours.

References

Anastasi, A. (1965). "Culture fair testing". Educ. Horizons, XLIII fall 1964.

Deutsch, M. (1973). "Happenings on the way back to the forum". Harvard Educational Review. 39, 523-547.

Ghuman, A.S. (1974). A cross cultural study of the thinking processes of English 'British' Punjabi and Indigineous Punjabi boys. Unpublished Ph.D. thesis, University of Birmingham.

Hanfmann, E. & Kasanin, J. (1937). "A method for studying concept formation". J. Psychol. 3, 521-40.

Jensen, A.R. (1973a) Educability and group differences. Methuen.

Jensen, A.R. (1973b) Educational differences. Methuen.

Jensen, A.R. (1972) Genetics and education. Methuen.

Jensen, A.R. (1969) "How much can we boost I.Q. and scholastic achievement?" Harvard Educational Review. 39, 1-123.

Jensen, A.R. (1973) "Level I and level II abilities in three ethnic groups". American Educ. Res. J. 4, 10, 263-76.

Jensen, A.R. (1968). "Patterns of mental ability and socioeconomic status". Proceedings National Academy of Science U.S.A.

Lesser, G., Fifer, G., Clerk, D.H. in Roberts (1965) "Mental abilities of children from different social class and cultural groups". Monographs of the society for research in child development. 30, 4.

Raven, J.C. (1969). Coloured progressive matrices. Lewis.

Raven, J.C. (1960). Standard progressive matrices. Lewis.

Renhard, D.H. (1971). An attempt to increase nonverbal intelligence test scores by the programming of the principles of inductive reasoning. Unpublished M.Ed. thesis, University of Birmingham.

Semeonoff, B., & Trist, E. (1958). Diagnostic performance tests. Tavistock.

Semeonoff, B. & Laird, A.J. (1952). "The Vigotsky test as a measure of intelligence". Birt. J. Psychol. 43, 94-102.

Semler, I.J. & Iscoe, I. (1963). "Comparative and developmental study of the learning abilities of negro and white children under four conditions". J.Educ. Psychol. 54, 38-44.

Stones, E. & Heslop, J.R. (1968). "The formation and extension of class concepts in primary school children". Brit. J. Educ. Psychol. 38, 3, 261-71.

Stones, E. (1970). "Verbal labelling and concept formation in primary school children". Brit. J. Educ. Psychol. 40, 3.

Vigotsky, L.F. (1962). Thought and language.

Wesman, A.G. (1968) "Intelligent testing". American Psychologist. 23, 4, 267-274.

PROBLEMS OF DESIGN IN RESEARCH ON TEACHING

Michael J. Dunkin.
Macquarie University.

Let me begin by making it clear that this paper is not about research on teaching, the profession, but that it is about research on teaching, the activity. Thus, nothing will be said concerning supply and demand of teachers, the processes by which they are socialized into the profession, the network of expectations they might encounter, how they came to choose teaching as a career, or the host of other important matters that could be discussed given the concept of teaching as a profession. Instead, the questions that underlie what I will be saying are of the type that Gage (1963) raised in the first Handbook of Research on Teaching:

1. What is the nature of teaching behaviour?
2. What are the determinants of teaching behaviour?
3. What are the effects of teaching behaviour?

Having reached that point, it is also necessary to place limitations upon the meaning of teaching behaviour. By teaching behaviour, I mean acts performed towards another or others in contexts where the expectation is that the latter will learn as a result. Note that this definition does not require that the performer of the acts be formally designated a teacher, that all the acts have behind them an intention directly to promote learning, or that learning, let alone "desired" learning, result from those acts. This definition conflicts strongly with the "success" notion of teaching according to which if there has been no learning, there has been no teaching, and departs somewhat from the "intent" notion of teaching in that it allows that some teaching acts might not have behind them an intention that learning result, thus avoiding the necessity to make inferences about teacher intent and distinctions between management and "genuine" teaching behaviour. (See Hyman 1972)

Note, too, that the definition does not require school or classroom contexts.

Teaching behaviour can occur between parent and child, doctor and patient, salesman and customer, bus conductor and passenger, and so on. However, for the purposes of this paper the classroom context will be assumed.

Given the above limitations, research on teaching in this paper refers to the study of teaching behaviour in classroom contexts. That is, I will be considering research that has involved the systematic observation and measurement of classroom teaching behaviour. This limitation means that I am excluding most of the research conducted over the decades on teacher effectiveness for a considerable proportion of that research has not involved measurement of teaching behaviour, and of attempts to "measure" teaching behaviour, many have not involved systematic observation of it.

To conclude this introduction, may I point out that while I have attempted to delimit the meanings of teaching, teaching behaviour and research on teaching, I will be using the term "design" in a very broad sense. Sometimes I will be using it to distinguish a long more conventional lines between experiments and correlational studies, but, in the main, I will be accepting a much wider range of decisions that researchers make as aspects of the design of their research.

THE NATURE OF TEACHING BEHAVIOUR

Now, to return to the three questions raised by Gage (1963). The first, concerning the nature of teaching behaviour, has received considerable attention in studies that have concentrated upon what have become known as process variables, and which I have referred to in the summary of this paper and elsewhere (Dunkin and Biddle, 1974) as studies exhibiting process design. These are natural history type studies that have focussed upon collecting instances of teaching behaviour through systematic observation, and classifying them according to some conceptual framework. Most of them have provided evidence that only partially answers Gage's first question. These have been attempts to test the validity of a set of concepts of behaviour in the classroom context. They tell us that such categories as conditional inferring, criticism, structuring and vagueness do occur, can be measured more or less reliably, and can be distinguished from one another by trained observers. Even if some of the specific categories have not stood up when evaluated by these criteria, we at least know with considerable confidence that several different, though broad, facets of behaviour occur and are observable. They have enabled us to frame questions about teaching in terms that have some empirical basis, and some less cautious souls have even taken evidence of the existence of these categories as sufficient for prescribing what teachers should do to enhance their effectiveness.

While some reviewers, for example Rosenshine (1970), have been concerned at the sheer number of observational instruments generated from process design studies, the fact is that many of them are fairly minor modifications of Flanders Interaction Analysis Categories. Thus, many theoretical constructs that are popular in educational literature remain to be translated into behavioural terms, and potential effects of and needs for innovations remain undiscovered.

Then, of course, there is still plenty of scope for the development of concepts that grow directly from instances of classroom interaction itself, through careful, laborious, and insightful scrutiny of the latter. A very good example of this is in the work of Professor A.A. Schorb, Dr. Brigitte Louis and their colleagues at the Institut für Unterrichtmitschau und Didaktische of the University of Munich.

THE DETERMINANTS OF TEACHING BEHAVIOUR

While solid progress has been made in developing valid concepts of teaching behaviour and in constructing instruments with which to observe and measure it, we know very little about teaching behaviours as variables. That is, we know little about their incidence, the conditions under which they vary and the shape of their distributions. And yet this type of information is crucial if we want to do research with them. For example, there is little support as yet for the assumption underlying many statistical procedures employed in research on teaching that teaching behaviour variables are normally distributed. Indeed, I have the impression from having reviewed some hundreds of these studies, that skewedness is more the rule than the exception. Many of the research reports, for example Wright and Nuthall (1970), include information that standard deviations approximate, even exceed, means. It is not clear whether these reflect sampling deficiencies or whether they are accurate indications of the actual distributions of the variables. However robust our statistical methods are of non-normality, proceeding on the assumption of normality could well result in our being denied important evidence concerning teaching. Thus, in my view, one of the greatest needs in research on teaching is for process designs to be implemented with carefully drawn samples of teaching behaviour with the express purpose of telling us more about the distributions of teaching behaviour variables.

It would be remiss of me not to mention in passing an elaboration of the process design. This is the process-process design where the concern is with exploring interrelationships among different variables or facets of teaching behaviour. It is not unusual to come across in the literature single variable prescriptions such as 'teachers should ask more higher-order questions" or 'teachers should increase their use of praise". But what are all the possible consequences of manipulating one or a few teaching behaviours? Some of them will probably be changes in other teaching behaviours. Let me give you an example from some research I have been engaged in over the last few years (Dunkin, 1975). It has involved a sample of lessons in social studies and mathematics at primary school level in the Western Highlands of Papua New Guinea. Among the variables I have been investigating are logical complexity and abstractness. Now some prescriptions for teaching behaviour involve increasing teachers' use of logically more complex, or if you like, higher-order questions. Others exhort teachers to enhance concept development by providing concrete rather than abstract experiences for children at certain stages of cognitive development. It just so happens that in the sixth grade classes I observed, in both subjects, there was a significant positive relationship between logical complexity and abstractness, such that as teachers ascended the logic ladder they also ascended the abstractness ladder. If I were to prescribe one of these ascents without training the teachers to turn the other into a descent, I might be responsible for counter-productive teaching. In short, process-process designs also offer us potentially valuable evidence of the complexity of teaching behaviour.

THE EFFECTS OF TEACHING BEHAVIOUR

At this point the abstract of this paper suggests that I should begin to talk about context designs in research on teaching. However, since a major part of context designs concerns research in relation to curriculum development, and since there were several papers on this very topic this morning I shall not say anything about them at this time.

Neither shall I say anything about presage designs at this stage, for I would like quickly to proceed to product designs where I think the most exciting developments have been taking place in recent years.

Product designs are those which attempt to answer the question, what are the effects of teaching behaviour? Researchers in this crucial area have, to my mind, had to face the most difficult problems there are in research on teaching. Several reviews of their efforts have been published in recent years, and all the reviewers have expressed strong dissatisfaction with the state of the art. Rosenshine and Furst (1971), after reviewing approximately fifty studies, concluded as follows:

> 'We note that the early expectations that the counting of relatively objective teaching behaviours would yield consistent, significant correlations with student achievement certainly have not yet been fulfilled.... After ten years of process-product research, fifty studies, and mixed results, some researchers would claim that such correlational research will not be productive in the future. Because of the limited research, and because of the methodological problems which may exist in most of these studies, any judgment on the worth of this research would be premature." (pp. 55-56).

Dunkin and Biddle (1974) concluded:

> 'In sum, most findings from this field must presently be presumed tentative: because we are not sure how strong they are, because we do not know whether they are independent of other effects, or because they have not yet been validated experimentally." (p. 359)

The above conclusions are gentle and optimistic compared with those of Heath and Nielson (1974) who, after re-examining the studies reviewed by Rosenshine and Furst, had this to say:

> 'Our analysis of this literature leads us to three conclusions:
> First, the research literature on the relation between teacher behaviour and student achievement does not offer an empirical basis for the prescription of teacher-training objectives.
>
> Second, this literature fails to provide such a basis, not because of minor flaws in the statistical analyses, but because of sterile operational definitions of both teaching and achievement, and because of fundamentally weak research designs.
>
> Last, given the well-documented, strong association between student achievement and variables such as socio-economic status and ethnic status, the effects of techniques of teaching on achievement (as these variables are defined in the PBTE research) are likely to be inherently trivial." (p. 481)

While there might well be agreement among reviewers on the first two conclusions reached by Heath and Nielson, the third one has provoked considerable discussion. Good, Biddle and Brophy (1975), for example, have devoted a whole book to justify the counter-assertion that "Teachers Make a Difference", while Nuthall was provoked to comment:

> "Those who have drawn the conclusion that teachers make no difference from the evidence reported by Coleman and others in Equality of Educational Opportunity (1966) have misunderstood the nature of the data analysis they employed in that study, and overlooked the ambiguity of the word 'difference'. What studies of the Coleman type may (or may not) have shown is that measured differences between teachers are not closely associated with measured differences between pupils. What such a finding might mean is very difficult to understand, but it certainly does not mean that teachers do not affect pupils and it does not indicate in any way how much they affect their pupils. It should not take large-scale research studies, or sophisticated regression analyses, to indicate that children who are taught to play the piano, do rather better at piano playing than children who have never received any lessons.
>
> We know, from substantial commonsense evidence, that teachers do have an effect. We know that in many clear-cut cases teaching does work. And we must know how teaching works before we can hope to be making (sic) anything more than accidental progress towards improving the quality of our schools. What we do not know is how teaching works". (Nuthall, 1974, pp.2-3.

No doubt Nuthall's comments are consoling to teachers and teacher educators who might otherwise conclude that they are inevitably wasting their time.

Nuthall goes on to address himself to the view that research on classroom behaviour can never be good research because it cannot meet the criteria of acceptable scientific procedure, and counters that argument by pointing out not only that traditional criteria are themselves open to question but also that attempts to make classroom research conform to them have resulted in the wrong questions being asked and the wrong kinds of answers being sought. His subsequent analysis of the situation leads him to conclude:

> 'Consequently, what it means to say that the present status of research on teaching is one of continuing production of weak and ambiguous results is that the data produced are not holding up very well within the traditional framework of experimental design and data analysis. To claim that the research is not worth pursuing because it cannot be made to fit the traditional criteria of good research design may be putting the cart before the horse. If data do not meet the assumptions involved in the traditional research designs, it does not follow that we should throw away the data." (p. 5)

One of the traditional assumptions underlying classical research designs is that events are likely to be associated with each other in a linear fashion, or that graphs of the relationships will approximate straight lines.

Instead, Nuthall demonstrates that for such variables as structuring before a question, structuring at the end of a question-answer type episode of discussion, and teacher criticism, the relationships with change in pupil achievement and attitudes might well be non-linear, with the optimal levels of performance varying from one context, such as fourth grade, to another, such as sixth grade. Obviously, this is a much more complex state of affairs than simple linear relationships. Furthermore, by re-analyzing data gathered by Wright and Nuthall (1970) in a way that permits discovery of non-linear relationships and avoids some assumptions underlying more conventional statistical methods, Nuthall demonstrated that the latter had produced quite misleading results, and that non-linear relationships with measures of pupil attainment were more typical and more likely to be statistically significant.

SUMMARY

In conclusion, may I be so bold as to summarize the type of arguments I have presented above in the form of recommendations for research aimed at enhancing our understanding of the nature of teaching and its effects:

1. We need to be much more flexible in the assumptions we make about the types of relationships that exist among process variables and between them and product variables, and in our approach to statistical methods. Nuthall's seminal paper has provided us with a very promising lead here;

2. We need to pay much closer attention to the influence of context variables upon process variables, relationships among them, and relationships between them and product variables. That is, we need replications of studies in one context in further studies where contexts are systematically varied;

3. In line again with Nuthall's suggestions, our attention might well be given much more to the question of how teaching works than it has in the past when emphasis was given to the degree to which teaching works;

4. We have not nearly exhausted the richness of the data available for analysis in teaching contexts. We need to continue the search for new concepts of classroom behaviour through both closer inspection of what is there, in the hope of insights occurring, and through attempts to identify the classroom behaviour equivalents of abstract concepts that abound in theoretical approaches.

5. Finally, we need to exercise much more care and imagination in our choice and measurement of product variables. We need not only to explore much more widely the range of possible outcomes of teaching in pupils, but also to design our tests of these to reflect the opportunities to learn that are provided in the lessons under observation. A very good example of the latter may be found in Hughes (1973).

REFERENCES

DUNKIN M J — A Study of Classroom Ineraction in Papua New Guinea Primary Schools. Interim Report. Macquarie University, 1975.

DUNKIN M J and BIDDLE B J — The Study of Teaching. New York: Holt, Rinehart and Winston, 1974.

FLANDERS N A — Analysing Teaching Behaviour. Reading, Mass: Addison-Wesley, 1970.

GAGE N L — Paradigms for Research on Teaching. In N L Gage (ed), Handbook of Research on Teaching. Chicago: Rand McNally, 1963.

GOOD T, BIDDLE B J and BROPHY J J — Teachers Make a Difference. New York: Holt, Rinehart and Winston, 1975.

HEATH R W and NIELSON M A — The Research Basis for Performance-Based Teacher Education: Review of Educational Research, 44, 4, 1974, 463-484.

HUGHES D C — An Experimental Investigation of the Effects of Pupil Responding and Teacher Reacting on Pupil Achievement. American Educational Research Journal, 10, 1, 1973, 21-37.

HYMAN R T — Contemporary Thought on Teaching. Englewood Cliffs, N.J.: Prentice-Hall, 1971.

NUTHALL G A — Is Classroom Interaction Research Worth the Effort Involved? New Zealand Journal of Educational Studies, 8, 1, 1974, 1-17.

ROSENSHINE B — Evaluation of Classroom Instruction. Review of Educational Research, 40, 1970, 279-300.

ROSENSHINE B and FURST Norma F — Research on Teacher Performance Criteria. In B O Smith (ed), Research in Teacher Education: A Symposium. Englewood Cliffs, N J: Prentice-Hall, 1971.

WRIGHT C J and NUTHALL G A — Relationships Between Teacher Behaviours and Pupil Achievement in Three Experimental Elementary Science Lessons. American Educational Research Journal, 7, 1970, 477-491.

POLICY AND PRIORITIES IN PUBLIC FUNDING OF EDUCATIONAL RESEARCH

POLICY AND PRIORITIES IN THE PUBLIC FUNDING OF EDUCATIONAL RESEARCH IN BRITAIN

John Nisbet.
University of Aberdeen.

My paper is essentially a discussion paper and as befits a discussion paper it contains questions, comments and possible answers. Four questions are explored. These are:-

1. Should there be (or can there be) a coordinated policy in educational research in Britain?
2. How do we establish a balance between central coordination and independent initiative?
3. Is it practicable (or wise) to try to establish priorities for educational research.
4. What styles or patterns of research organisation and funding are best?

1. A COORDINATED POLICY?

Should there be (or can there be) a coordinated policy in educational research in Britain? If so, how do we set about it? What are the first steps towards a greater degree of coordination? What are the weaknesses in the present division of responsibility (informal and unwritten), which might crudely be summarised as: DES/SED policy-oriented, SC curriculum and examinations, ERB theoretical, UGC providing a research floor and the foundations as free-lance.

The idea of a coordinating Educational Research Council was widely debated ten years ago:

> "While the bases of these developments in research and development work were being laid in the Ministry (later the Department) of Education, in the curriculum development group and in the independent Foundations, arguments were going on about the possibility of establishing an Educational Research Council which could support both research and the training of educational research workers. Many educationists wanted such a council to be a separate body, parallel to the research councils in other fields. Many social scientists argued for the inclusion of education within a Social Science Research Council. The question was remitted to the Heyworth Committee; the social scientists' views prevailed, and, once established, the Social Science Research Council set up an Educational Research Board under the initial chairmanship of Lord James, parallel to its other committees for sociology, economics, psychology and the rest of the social sciences." (Taylor, op. cit., pp. 20-21).

Some of those who had high hopes when the ERB was set up have been disappointed. At a meeting called by the Universities Council on the Education of Teachers in January 1972, the view was expressed that "the ERB had been inhibited from taking what could be a vital role by its location within SSRC". Taylor (Research Perspectives in Education,) Routledge, 1973 discusses the SSRC and concludes: "The scale and nature of its present activities in the field of education are not such as to ensure that it can play the major role that the advocates of an independent Educational Research Council had seen as necessary" (p. 25).

The pressure for a separate Council has lessened, but the problem of coordination is still seen as urgent. In 1971, a UCET policy document expressed the point in a moderate and restrained statement:

> "There should be consideration at national level of the possibility of establishing better machinery than at present for the identification and discussion of research priorities and the coordination of research policies and initiatives. UCET is of the opinion that information coordination of the kind supplied by the existing pattern of overlapping membership of decision-making bodies, committee assessorships, personal contacts at conferences and meetings, and ad hoc discussions between directors and staffs of research funding agencies is inadequate to meet the needs of the next decade. The opposing dangers of creating a monolithic research structure and diverting resources from research itself to the essentially unproductive support of coordinating and policy formulating activities must equally be avoided."

Communication among the research-funding bodies has improved greatly in recent years, and ERB continues to press for closer liaison. In some cases, the need is for collaboration or integration of projects; in others, demarcation; in others, consultation and exchange of information. But probably a majority of researchers would agree with the view that:

> "Coordination is useful, provided it does not throttle intellectual independence and initiative. In the complexities of the social sciences, complete coordination of research would require omniscience and should never be attempted."
> (Conrad, Educational Research, 196[illegible])

2. CENTRAL COORDINATION OR INDEPENDENT INITIATIVE?

How do we establish an appropriate balance between central coordination and independent initiative, between the responsive and the initiating roles? Should there be closer monitoring of some publicly funded research projects, with other projects being designated as free-ranging, or high-risk? Does monitoring mean interference, and is interference necessarily bad?

Since 1970, the pendulum has swung towards greater control over research, more initiatives from the centre, more commissioning of research and closer monitoring of projects. In 1970, Mrs. Thatcher said:

> "There was clearly only one direction that the Department's research policy could sensibly take. It had to move from a basis of patronage - the rather passive support of ideas which were essentially other people's, related to problems which were often of other people's choosing - to a basis of commission. This meant the active initiation of work by the Department on problems of its own choosing, within a procedure and timetable which were relevant to its needs."

In 1971, Rothschild in the Government Green Paper, A Framework for Government Research and Development (HMSO) was convinced that the solution was simple:

"This report is based on the principle that applied R&D....must be done on a customer-contractor basis. The customer says what he wants; the contractor does it (if he can); and the customer pays."

In July I972 the Government White Paper on Research and Development accepted the principle:

"Departments, as customers, define their requirements; contractors advise on the feasibility of meeting them and undertake the work; and the arrangements between them must be such as to ensure that the objectives remain attainable within reasonable cost. This is the customer/contractor approach. The Government reaffirms its intention, announced in the Green Paper, of extending it to all its applied research and development."

Perhaps in I97I it was necessary that this over-simple point should be made crudely, to inject a sense of reality into the sometimes airy-fairy world of social science. But it is surely time for a more intelligent and sensitive statement to describe more accurately the complicated procedure of commissioning and funding projects which has developed. In several areas of research, initiatives from the centre have proved difficult and slow, requiring protracted dialogue and negotiation between so-called customer and contractor. A range of different solutions has been adopted: for example, the SSRC programme on Transmitted Deprivation, the ERB pre-school research programme, NDPCAL procedure as outlined in the I975 report, Two Years On, are all different in style. ERB initiatives have taken a variety of forms; seminars have proved one of the most successful procedures. Whether the initiative comes from researcher or the central organisation, there has to be a period of discussion and negotiation before the design and scale of the research is settled and the grant is made. The ERB does this for programmes and large projects, but not for medium or small grants; other organisations do it for all projects.

For their part, researchers must be responsive to the interests (both short-term and long-term) of those who provide facilities for their work, and must accept that they are accountable for what they do with other people's money. Monitoring by a steering committee, and the submission of regular progress reports are conventional checks which the research community accepts, but they are often ineffectual. NDPCAL is using "stepped funding": projects are funded in steps of one or two years, with subsequent steps dependent on successful "mid-term" evaluation. We need to explore styles of monitoring to establish the principles which determine the appropriateness of different styles to different situations. Rothschild is wrong in treating the matter as a confrontation of two groups of interests, customer and contractor. There are three groups involved: the sponsor (who has funds), the expert (who has research skills) and the user (who has problems). If two of these team up, they can force the third into a subordinate role.

3. PRIORITIES FOR EDUCATIONAL RESEARCH?

Is it practicable (or wise) to try to establish priorities for educational research? By the time priorities are identified and agreed, they may be out of date. Is it possible to identify emerging priorities for encouragement, and overworked areas of research where discouragement is required?

If it is practicable, it certainly is not easy. Many lists of priorities for educational research have been constructed - in Britain, USA, Canada, Australia. They are all different, and they are short-lived. Radford (Research into Education in Australia I972, AACRDE, Canberra I973) is opposed to the laying down of priorities in research by a central body:

"My reason is simple. Such a laying down of priorities to me implies an impossible omniscience, and lays up trouble for itself.... Provided that those engaged in research develop adequate channels of communication between themselves.... I believe there will not be any greater gap between the nature of problems and the information from research available to solve them, than there would be were there to be a central determination of a limited number of priority areas in which a lone study would be supported. It is well to remember that, not very many years ago, "education as investment" and "manpower studies" loomed very much larger as matters of research than now seems warranted by later experience.... Had the major part of the research apparatus swung over to such studies in I965 or I966, a good deal of work now known to be more valuable would not have been done."

He argues that "the common sense and perceptiveness of the workers should be backed, in the belief that no area will suffer." Sceptics advance the opposite views:

"An R&D program left to a self-regulated system determined by the whims of researchers inevitably leads to gaps, undesirable duplication and wrong priorities." (Lamontagne, Annual Canadian Conference on Educational Research, I97I).

"The country's needs are not so trivial as to be left to the mercies of a form of scientific roullette, with many more than the conventional 37 numbers on which the ball may land". (Rothschild, I97I).

Those with an academic background and an awareness of the history of science are likely to doubt whether a country's needs can be set out in a simple list of priorities. The SSRC report for I970-7I (written before Rothschild stated:

"It is not so much a matter of an ordered hierarchy of priorities, as a process of grasping at opportunities presented by an almost accidental coagulation of interest among a group of able research workers around a chosen problem in order to shift a fronier of knowledge forward".

"Almost accidental" means that even on this extreme view there are priorities. Indeed, there are priorities implicit in every decision that is made; one does not escape priorities by refusing to declare them.

Priority lists may be useful aids if they are not applied rigidly. They may be loosely defined. accepted as general guides only, or allowed only limited influence on decisions. The researcher should have some provision to allow him to follow up what he thinks is important - even Rothschild allowed IO% of research funds for this. Periodic reviews can make explicit the priorities which have operated. Thus Ward (Resources For Educational Research and Development NFER I973) calculated the distribution of research resources in education. Pre-school and primary research received 0.9 and 2.4% respectively of research funds, higher education took I6%, secondary took nearly 50%, and further education was the Cinderella with 3.7%. Is this the right distribution or does it reflect public pressure or professional power, or just chance? In the I968-70 period, Ward reckoned that less than IO% of funds were devoted to basic research, as opposed to development work.

"There is no coherent systematic rationale for the allocation of funds in....the knowledge base on which such decisions are made is weak and hence it remains ad hoc and partial".

Ward's data are already out of date and perhaps BERA should consider making

"a study which would show, at least in broad terms, types of research in education at present being undertaken throughout the country, areas or problems that seem to be incidentally attended to, conditions under which research is normally conducted in this country, and what educationists who are involved in research themselves believe to be the areas most in need of stimulation and support.. ..the number and types of researchers at work throughout the country; the forms of training that now exist for producing competent research personnel, and the adequacy of existing oportunities for recruiting and training research workers." From the Report of the Australian Advisory Committee for Research and Development, set up in I970).

This paper has focused on three questions:

I. a coordinated policy for educational research;

2. the balance between the responsive and the commissioning styles of funding;

3. priorities in educational research.

A connecting thread is the question, how effectively can educational research be planned and organised without imposing such limits on freedom of thought and action that the research itself becomes limited in its scope and applicability?

Depending on the answer to these basic questions, other groups of questions arise - for example, on the appropriateness of the existing machinery for funding:

4. WHAT FORMS OF FUNDING ARE BEST?

What styles or patterns of research organisation and funding are best? (Best for what?) Is the 'project' the correct unit (usually 2-3 people, 2-3 years, building up rapidly, running down rapidly, fixed-term appointments, no career structure)? Should we move towards concentration of effort, aiming to develop centres of excellence and semi-permanent educational research units? Perhaps locally based regional research (or R&D) units would be better. Should the main effort be put into improving the provision for training research workers, or winning public confidence in research, or theoretical studies, or improving the infrastructure of research - or a mixture of these?

A basic reference.
The topic is concisely reviewed by William Taylor in the chapter in H.J. Butcher and H.B. Pont, Educational Research in Britain 3 (ULP 1973):

pp. 19-21 increase in research expenditure (ten-fold in six years 1962-68); five main sources DES/SED, Schools Council, ERB/SSRC, independent foundations and UGC);

pp. 22-27 present position issues arising;
pp. 28-30 future prospects; need for coordination.

USING CLASSROOM OBSERVATION IN CURRICULUM INNOVATION

THE CONTRIBUTION TO CURRICULUM INNOVATION OF CLASSROOM OBSERVATION BY TEACHERS

Wynne Harlen.
University of Reading.

To date there has been a tendency for curriculum innovation to be too narrowly conceived as concerned primarily with the production of new materials and resources for teaching. A decade of curriculum projects geared to the development of tangible products, with activity ceasing immediately after, if not before, publication has resulted in much less real change in schools than might have been expected and thought to be potentially possible. Lately it has been acknowledged that the process of curriculum innovation is concerned only in part with devising new programmes or resources and that the provision of guidance and support for teachers attempting to use the products is of equal importance if change is to take place and be lasting. The hitherto neglected part of innovation, which requires changes in teachers' attitudes and values as well as in their practices, is now receiving some attention in the dissemination programmes built into project proposals.

The relative neglect of preparing teachers as compared with preparing materials is particularly serious in this country where curriculum projects have generally tried to preserve teachers' freedom of action, and rightly so. But such freedom, without guidance, can result in the all too familiar situation in which new materials are used in old ways. Projects have provided for teacher choice in using materials so that activities may be selected or adapted according to the experience and learning characteristics of particular pupils, i.e. there is room to try to match activities to pupils rather than vice versa.

However if these intentions are to be translated into practice teachers must have information about their pupils in order to make decisions about suitable activities. The information a teacher requires will concern those personal characteristics, attitudes and preferences which affect a pupil's learning, as well as the previous knowledge and level of development of mental and physical skills. This information is needed for on-going decisions, and for many reasons cannot be provided by conventional methods of assessment. Observation, however, is a method which can be readily employed by teachers for gathering the feed-back required, and this is an important use of classroom observation in a new context - in the hands of teachers and not of outsiders to the classroom.

Clearly to be a useful method for collecting information observation cannot remain at the level of general impressions and subjective judgements. It would be particularly unhelpful in the context of innovation for observation to be unguided, for there would be a danger of the observations being influenced and channelled by the structure of outdated goals and discarded values. A structure for observation is required, which comes from considering the directions of the pupils' learning and the characteristics which affect the learning. Within this framework teachers' observations can be turned to pupils' behaviours relevant to the goals and values of the innovation; thus classroom observation by teachers has an important role in reinforcing their appreciation of new goals and processes.

The work of the Schools Council Project Progress in Learning Science exemplified the points which have been made. It originated in response to the difficulties which prevent many teachers in primary and middle schools from incorporating science into their pupils' work at all effectively, despite the large efforts which have been put into curriculum development in this area. The project has produced check-lists to provide a framework for teachers' observations of their pupils. Each list covers about twenty-five abilities, concepts and attitudes which are considered the more important goals of science activities, and statements in the list indicate typical behaviours at three levels in the development of each skill and concept. Other products of the project aim to help teachers make decisions about appropriate activities at different stages and suggest the approaches which foster pupils' positive attitudes to their work and to themselves. These materials are introduced to teachers in the context of group or individual study of the ideas behind them, of discussion of the meaning of the goals and statements in the check-lists, and of experiences which enable teachers to develop the skills of observation and of interpretation of children's behaviours.

APPENDIX

Examples from the check-lists of the Progress in Learning Science Project.

Items included in the check-list for Early development.

Observing	Curiosity
Raising questions	Originality
Exploring	Perseverance
Problem Solving	Willingness to Cooperate
Interpreting Findings	Open-mindedness
Communication Verbally	Self-criticism
Communicating Non-verbally	Responsibility
Applying Learning	Independence
The Concept of Causality	
Classification	
The Concept of Time	
The Concept of Weight	
The Concept of Length	
The Concept of Area	
The Concept of Volume	
The Concept of Life Cycle.	

Examples of statements in the check-list

Problem solving

1. Generally unable to approach a problem without help, or gives up if his first idea does not succeed.

2. Tries one or more ways of tackling a problem without much forethought as to which is likely to be relevant.

3. Identifies the various steps which have to be taken and tries to work through them systematically.

Open-mindedness

O. Tends to stick to preconceived ideas ignorning contrary evidence and behaving as if unaware of the existence of opinions or findings different from his own.

1. Changes from idea or opinion to another inconsistently, being influenced by the authority behind alternative views rather than the strength of the evidence or argument.

2. Generally listens to and considers all points of view and relevant evidence; accepts ideas different from his own if the evidence is convincing.

OBSERVATION OF NON–VERBAL ASPECTS OF TEACHING BEHAVIOUR: IMPLICATIONS FOR TEACHER EDUCATION CURRICULA

Peter Chambers.
West Midlands College of Education.

INTRODUCTION

Like most decisions about what should constitute the appropriate curriculum for a teacher in training, the decision to give this paper is based on hunch and speculation. There is so little in the way of obvious rationales for developing Teacher Education curricula that further speculation need not be harmful and a description of the processes involved might even be positively helpful. Some years ago an ATCDE Research Committee's Working Party initiated a paper on the objectives of Teacher Education, which worked its way through CRITE and the SSRC seminars. It placed the emphasis clearly on improving the classroom performance of a "beginning teacher". As such, this was quite uncontentious, but where few of us could agree, was in identifying, with any precision, the elements in that performance which could be linked to training. Nevertheless, I used that paper as a working paper for developing my Department's new syllabuses and at the same time (but quite coincidentally) found a new slant in my own researches into non-verbal communication in small group teaching situations. This suggested to me a starting point for curriculum development in Teacher Education. What I was trying to establish was what set of social constructs based on systematic observation of teacher/pupil behaviour in classrooms could help the development of curricula in Teacher Education. This is, of course, far too ambitious a project for the single research worker, but, if curriculum developers can be involved, directly or indirectly, in examining classroom interaction, even as individuals, there are the makings of a legitimate rationale for Teacher Education Curricula innovation. This paper is offered as a case study of one such research worker cum curriculum planner.

NON-VERBAL COMMUNICATION

Ironically, it was when my own initial research hypothesis was proving untenable, that the most useful theory emerged. My original interest in non-verbal communication had arisen out of studies of sensitivity training with students (Chambers, 1969 and 1973) and led me through Goffman to the work of Galloway (1968). My intention was to show that the non-verbal cues that reveal personal styles of teaching could be measured systematically and could be linked to the different teaching tasks that make up a teacher's activities. Pilot studies were undertaken to refine an instrument for measuring a teacher's non-verbal "style" of teaching. Following the work of Galloway (1962), who had developed a matrix of non-verbal cues in pupil-teacher interaction by extending Flander's categories of interaction analysis (Amidon and Flanders, 1963), a multiple matrix has been devised which had been systematically tested in primary schools. The main dimension is concerned with domination/integration and so an attitude scale was constructed to measure the degree of authoritarianism revealed by the teachers. This followed conventional Thurston and Chave lines (1929) (Edwards, 1957). The Scale was initially tested against people's perceptions of authoritarianism and achieved very high correlations with both expert opinions and pupil perceptions. A small sample of scores was also compared with scores from the same subjects on the Rokeach Dogmatism scale and although the results are positive, the relationship between dogmatism and authoritarianism is sufficiently problematic to reject further exploration of this aspect.

The main pilot study consisted of observing teachers in a variety of primary schools all undertaking small group teaching. Two main data collecting sessions were undertaken. All the teachers completed the attitude scale and were observed by a minimum of two observers, together with close circuit television cameras, for periods of up to five minutes. During this time the observers scored the teacher's non-verbal behaviour on the matrix proforma. Later, they watched the television clips and scored the teacher's behaviour afresh. After initial training periods using television film, there has been a high measure of inter-observer reliability and, as might be expected, almost perfect correlation between scores made on "live" performances by teachers and on television recordings. Since there can be nothing standardised about the matrix scores and those from the attitude scales, non-parametric statistical techniques have been used throughout and their shortcomings are readily acknowledged. Nevertheless, it was worth noting the high consistency between scores given by different observers (C = O.79 and O.85). Comparison of these non-verbal scores with the Attitude Scale Scores using Spearman's Rank Difference Method for calculating correlations of coefficient also produced high positive correlations. Within each population, there were higher correlations than for the total population, but the pilot study was felt to be sufficiently encouraging to continue the experiment.

Population A	13 teachers	O.93
Population B	2O teachers	O.75
Populations A and B combined	33 teachers	O.64

TEACHER EFFECTIVENESS

The main study is now under way in a further series of primary schools and confirms a process that the combined scores suggest. The correlations certainly exist, but it is the differences that are more important than the similarities.

The first population consisted entirely of genuine volunteers, who were, by and large, recognised as very able teachers. The second population also consisted of 'volunteers', but some were more willing than others, and there were three teachers who were considered to be little more than competent by their head teachers. Two of these three showed very large discrepancies between their "non-verbal scores" and their authoritarian attitude scale scores. Since this information was gathered after the event, it was only used to guide the next stage of the enquiry.

It was therefore decided to get an independent estimate of the degree of effectiveness of each teacher and to compare these scores with the degree of discrepancy between the teaching "style", as revealed by the non-verbal cues on the matrix pro forma and the "expressed attitudes" as measured by the authoritarianism scale. The estimate was gained, in the first instance, from the teacher's head teacher; from an experienced college tutor viewing the television clips; and from an independent qualified teacher, using a pro forma based on a College's evaluation schedule.
In this way it was hoped to test the hypothesis that dissonance between expressed attitudes and non-verbal teaching style would be an indication of an ineffective teacher.

Twenty teachers were filmed in story telling sessions in an attempt to reduce the number of tasks involved and also because this was thought to be a useful initial task for student teachers.

The interim results are as follows:

Teacher	Scorer	NVS	Auth	Eff
1	A	186	100	16.34
	B	185		
2	A	152	46	17
	B	140		
3*	A	143	41	17
	B	122		
4	A	98	60	10.67
	B	94		
5	A	104	24	14.67
	B	102		
6	A	150	26	13.34
	B	148		
7	A	50	63	8
	B	60		
8	A	134	17	12
	B	134		
9	A	141	-6	12.67
	B	139		
10	A	92	24	11.34
	B	104		
11	A	157	81	18.67
	B	161		
12	A	118	65	10.67
	B	113		
13	A	59	50	10
	B	77		
14	A	159	117	19
	B	168		
15	A	125	-1	15
	B	133		
16*	A	157	42	18
	B	136		
17	A	115	-43	10.5
	B	101		
18*	A	153	30	17
	B	122		
19*	A	160	35	18.5
	B	138		
20*	A	64	70	7.5
	B	92		

Inspection reveals very close correspondence between the observers and this is supported by the calculations of the coefficients of correlation, which are O.90. Those teachers marked with an asterisk represent the most extreme differences and re-analysis after re-running the video-tape made the judgments very much closer. Comparisons were made between non-verbal scores and authoritarianism scores and between non-verbal scores and teacher effectiveness ratings.

These correlations are as follows:

Non-verbal and authoriarian O.15; non verbal and effectiveness O.83; authoritarianism and effectiveness O.14.

The relationship between non-verbal scores and teacher effectiveness is clearly revealed by these figures, but the hypothesis about dissonance is further supported by analysis of the differences between scores on the Non-verbal Measure and the Authoritarianism Scale Scores. The correlation of coefficient between effectiveness ratings and ranked differences of NVS and Authoritarianism scores is O.685. The mean differences between different effectiveness ratings appear to confirm this.

Teachers Rating	N	Differences in Rankings on NVS Scale and Auth. Scale	Mean Difference
A	3	1, O, 7.5	2.80
B	7	1, 4, 10, 4, O, 6, O.5	3.64
C	8	9.5, 11, 7, 8, 6, O.5, 10, 11	7.88
D	2	14, 14	14.00

The data drawn from this small population would therefore suggest that there is a clear relationship between the teacher's non-verbal communication and his or her effectiveness in teaching through the medium of story. Additionally, a clue to this "ineffectiveness" can be found in the dissonance between a teacher's teaching "style", as measured by an indirect scale of authoritarianism, and the impressions he or she "gives off" in his or her non-verbal communcation.

Further, the high correlation between the non-verbal Scores and the effectiveness rating is actually an indication of a close relationship between non-authoritarianism as revealed in non-verbal cues and teacher effectiveness since the non-verbal scores are orientated this way on the dominant/integrative dimension.

Although this paper does not permit a full discussion of this issue, it seems clear that the favoured role model among the experts consulted in this enquiry, and in the author's perceptions (the critical variable?) is of the integrative, child-centred variety. The numbers of effective authoritarian teachers are minimal. Using the evidence of the attitude scale, it seems that the teachers clinging to the old role model of authoritarian teaching are perceived as the ineffective ones.

These very tentative figures do, in fact, support impressions based both on general observations over a period of years and also the very full and detailed qualitative observation that lies behind the small scale statistical quantification of these experimental procedures, which manifestly fail to do justice to the richness of the phenomena observed. Both kinds of data suggest, firstly, that meta-incongruence between teaching intention and teaching performance as revealed to an audience in non-verbal communication is an important clue to what makes some teachers ineffective and, secondly, that this relationship between intention and performance is the product of the sort of role models, popularized both by training procedures and current fashions in schools.

SOME IMPLICATIONS FOR TEACHER EDUCATION CURRICULA.

There are three main aspects of this kind of research that can be linked directly to the development of curricula in Teacher Education. The first plainly concerns the place of classroom experience and the use to which it is put in training teachers.

The present system of "**Block** Practices", be it consecutive or concurrent, would seem to lack the kind of precise targets that this enquiry suggests exist. As it is, the concept of School Experience would appear to be too global, too dependent on serendipitous happenings, to be justified if we can pin-point something as crucial as the way a student teacher presents his "front" in the classroom. If impression-management is an important as this small scale study suggests, then there is a strong case for making initial experiences relate to the particular aspect of what sort of impressions are given off. We have developed some sporadic approaches to micro-teaching in College and in the case of two first year groups have tried to concentrate on non-verbal communication. We have also tried, but not continued with, a "group practice" experience in schools in which the CCTV cameras filmed students' interventions into children's learning. The same students also completed the attitude scale and we discussed both results with them.

Unfortunately, a CNAA validation visit overtook our attempts to build the experience into our curriculum. Nevertheless, our experience would support attempts to build quite specific school experiences into the College curriculum that set out to teach the skills of impression management for students.

The analysis of teaching, using schedules like the California Stanford Teaching Appraisal Guide or techniques like Interaction Analysis, would fall into this category. The perceptual load that faces the student teacher when he first enters a classroom would seem too enormous to rely on a global approach and the use of techniques that focus perception both on the classroom activity and, even more importantly, on the student teacher's involvement and performance in it would seem immeasurably superior. Further, the curriculum planner's own attempts to measure such activities certainly add a crucial dimension to the way the experience is used.

The second aspect of curriculum development concerns the way such classroom experience, or simulated experience, or self-analysis procedures, are related to the total college curriculum. The present approaches are by and large compartmentalized and although there are considerable pressures today to make curricula coherent and inter-connected (Chambers, 1975), even the best expect the integration to occur in the student's mind. The use of the attitude scale has brought home forcibly how little attention is paid to the way students build up their role models of acceptable and effective teachers. The use of small scale teaching situations, their analysis and the relationship of such analysis to the indirect measurement of a student's attitudes to teaching offers a central thread around which training can be based. It obliges students to define their own intentions, and therefore their role models, and to see how these become manifest in their own teaching and, if the curriculum is sufficiently thorough and rigorous, to relate this to their own potential effectiveness as teachers. Certainly, in the early stages of the student's course, it offers a sound conceptual framework on which each student can plan his course. On the basis of consistent role models, the student can see more clearly what he may need in the way of Subject material, of Educational Theory and of Professional Courses. Here would be an admirable place to consider Role Theory and the implications models have for action in the classroom. It is more logical though to consider the data as the basis for individual decision-making by any given student or group of students about their courses. It is perhaps asking a lot of the various vested interests that construct present college curricula to allow the student to develop his course in the light of specific analysis of his early teaching experiences, but these tentative figures do support such a case.

Finally, there is a diagnostic element suggested by these figures that can affect the selection of students as well as the shaping of their curricula. With the increasing trend towards the diversification of College courses, selection for teaching becomes a very real possibility. The use of story-telling in the second place of this enquiry was a deliberate attempt to see whether a relatively easily controlled activity, eminently suitable as an initial teaching exercise, could be used to identify the wide range of non-verbal activity that makes up a teacher's repertoire. The results so far suggest that it can. The analysis of the non-verbal elements revealed and their comparison with the attitude scale scores appears to offer a promising way of identifying the potentially ineffective student at a very early stage in his career. Such diagnosis can provide an additional factor in planning individual student curricula quite apart from any spin-off in vocational guidance (or non-selection). It can provide clues about remedial programmes, for example, short-cutting procedures such as those practised at Coventry College in deciding who are the weak candidates for their experimental micro-teaching programme.

Further suggestions are left to others. As it is, the degree of inference from those very limited and imperfect data is already great. There is only a small gap between information-supported theorizing and special pleading. Nevertheless, it is hoped that this account of how one man's involvement in research based on classroom observation can influence innovation in the curricula of Teacher Education. It is hoped it will encourage others not only to do the same, but to do it infinitely better and with more courage and conviction.

REFERENCES

Amidon, E., and Flanders, N.A., (1963), The Role of the Teacher in the Classroom. Minneapolis: Amidon.

Chambers, Peter, (1969), "Training for Social Skills: A Preliminary Enquiry into Methods of Measuring Social Awareness," Paedagogica Europaea, Vol. V, 1969.

Chambers, Peter, (1973), "An Inter-disciplinary Enquiry into Group Processes and its Implications for the Education of Teachers," Creative Drama, Vol. 4, No. 5.

Chambers, Peter, (1975), "Course Validation and Curriculum Innovation," Education for Teaching, Summer 1975, No. 97.

Edwards, Alan L., (1957), Techniques of Attitude Scale Construction. New York: Appleton-Century-Crofts.

Galloway, Charles, (1962), An Exploratory Study of Observational Procedures for Determining Teacher Non-Verbal Communication, Unpublished D.Ed., dissertation, University of Florida.

Galloway, Charles, (1968), "Non-Verbal Communication," Theory into Practice, Vol, VII, No. 5, Columbus: Ohio.

Goffman, Erving, (1969), The Presentation of Self in Everyday Life. Harmondsworth: Penguin.

CLASSROOM OBSERVATION USED IN CURRICULUM EVALUATION

Sally Brown.
University of Stirling.

PURPOSES AND SETTING

There are at least three different levels at which the success of curricular innovations can be judged:

1. at the lowest level, teachers or administrators may or may not accept the innovation and attempt to implement it;

2. at a second level, the intended innovation may or may not be adequately implemented: teachers may not understand what was being prescribed by the curriculum developers or they may find it impossible to do what was intended.

3. at a third level, the innovation may be satisfactorily implemented but may or may not lead to the intended improvements in pupils' learning.

Our classroom observation study at Stirling is being used to examine the 'success' at levels 2 and 3, of a number of innovations considered as important in the implementation of a curriculum (Scottish Integrated Science, Scottish Education Department 1969). The focus is on specific aspects of teachers' and pupils' behaviour related to the particular innovations and our tasks are:

1. to describe quantitatively these aspects of teachers' behaviour;

2. to identify the extent to which and the ways in which these behaviours are related to pupils' achievement of the course's objectives (knowledge and understanding, attitudes and, perhaps, laboratory skills).

The innovations selected for study are: guided-discovery learning, integration of subject disciplines, teaching a common course to mixed ability groups, and teaching towards objectives.

Our exploratory study revealed that science teachers felt they had insufficient information on, and little understanding of, the four innovations (with the possible exception of the last one). The innovations are of a general nature and interest, but the classes we are studying are all mixed ability, mixed sex, taught all their science by one teacher, in the first year of comprehensive education, in schools large enough to have three science departments, and the content covered consists of two specified sections of the Integrated Science course.

There are, of course, differences among the schools in architectural characteristics, resource constraints and organisational characteristics of science departments. We have a further strand of the study which is exploring what ways, if any, these various departmental contexts, within which science teachers in different schools work, influence the aspects of classroom behaviour with which we are concerned.

The descriptions of teachers' behaviour will examine differences among schools and differences among teachers trained in different subject areas (biology, chemistry, physics), forming a quasi-experimental study with the teachers' behaviour as the dependent variable. In the process-product study, however, the dependent variables will relate to pupils' achievement of course objectives, and particular attention will be paid to the achievement of sub-groups of various abilities within classes.

CHOICE OF PROCESS VARIABLES.

We are not necessarily monitoring the teachers' behaviour in relation to the intentions of the curriculum developers since those intentions are not made sufficiently clear by the curriculum document. For example, guided-discovery is described as some sort of compromise between the approaches to learning of Bruner and Skinner; integration is taken at different points to mean the 'unity of all knowledge', 'the unity of science' 'interdisciplinary study', and 'one science teacher per class'. The variables we will be concerned with will a) reflect the various meanings that may be ascribed to the four innovations, b) be chosen such that classroom events may be categorized in terms of them, c) correspond as closely as possible to the ways in which teachers tend to describe their lessons, and d) reflect the pupils' opportunities to learn the criterion material (including reading or worksheet assignments, and exposure to AVA material) and to practice skills relating to the curriculum objectives. These variables are not, therefore, derived from any specified established theory except insofar as the curriculum developers' hypotheses that certain behaviours (such as those related to guided-discovery) would lead to certain outcomes are themselves based on theory (but this they have not made explicit).

Our sources of low-inference variables are discussed under 'The Classroom Instrument'. We are still hoping to use pupil questionnaires as a source of high-inference variables, probably along the lines of the Harvard Project work with the 'Learning Environment Inventory'. However, our preliminary questionnaire, concerned with the pupils' perceptions of the goal direction, difficulty, speed, differential attention given to pupils, pleasantness of environment, and disorganisation of their science class, was acceptable to some teachers but not to others.

THE CLASSROOM OBSERVATION INSTRUMENT

Audio recordings of 6 double periods (70 to 80 minutes) with each teacher will be made using a stereo tape recorder. One tape track will record from a radio-microphone on the teacher and a second track from an observer's microphone. The observer will record general aspects of teacher activity, pupil activity, and grouping of pupils, and the behaviour of individual pupils of various predetermined abilities during practical work; in addition speakers and groups/individuals with whom the teacher is interacting will be recorded.

UNIT	CHARACTERISTICS
Segment cf Gump 1967 'contained segment'	New segment identified by change in teacher activity, pupil activity or pupil grouping. Teacher activity: 'not-involved', 'watching-helping', 'demonstrating', 'discussing', 'instructing', 'attending to other source', and 'directing action'. Pupil activity: 'not involved','attending', 'consulting references','writing' 'discussing', 'working with scientific materials' and'readying'. Pupil grouping: 'whole class','groups/same task','groups/different tasks', 'individual work/same task' and 'individual work/different tasks'.
Transaction related to Topic (cf Sinclair and Goulthard, 1975, Gallagher, 1970, Bellack et al, 1966	New transaction identified by change in topic, segment, or interacting group. Topics: 'curriculum-substantive'*, 'other science substantive', 'routine procedural', 'disciplinary', and 'social'. *Curriculum substantive classified by content, level of instructional intent and level of abstraction.
Move (cf Bellack et al, 1966 and Coulthard 1975.	New move identified by change of speaker, or change in function of talk of single speaker. Types of move: 'framing' 'focussing', 'informing', 'directing', 'questioning' 'responding', and'reacting'.

Table 1 shows the units to be used to identify and categories the classroom events. These units were chosen with the following in mind:

1. A fairly large unit (segment) would be needed to describe the general strategies of the teacher and the context within which various tactical behaviours occur.

2. If we are to hypothesise ways in which teachers might teach more effectively, the units must be comprehensible to teachers and as close as possible to the ways in which they conceptualise the teaching.

3. The four innovations under study pointed to a system that allows for categorization of teachers' and pupils' types of activity, science content of the course, the tasks set and the interacting teacher and pupil roles related to these tasks.

4. It seems sensible to make use of other developed systems of units where these are appropriate.

The 'segment', as the largest unit, provides a coherent description of regular features of the science class such as changes in pupils' activity, or grouping of pupils and tasks, or role played by the teacher.

The 'transactions' contained within segments are discussions or other activities concerned with 'topics'. The pedagogical unit here is the 'topic' since the 'transaction' boundaries are identified largely by linguistic features. The 'topic' has not, however, been used as the analytic unit since we required a unit that would change with change of 'segment' or change of interacting groups (e.g. teacher moving round from group to group during pupils' practical laboratory work), and it is likely that the same topics would continue across such boundaries.

The control of the context of the teaching to two specified sections of the course provides the opportunity to prespecify and categorize most of the substantive topics. To this end a content analysis will be made of the specific objectives and syllabus content, the appropriate chapters from the textbook, and the 3 types of worksheet used by the schools for these two sections. The same sources have been used in the construction of criterion measures of pupil achievement, thus enabling us to relate pupil performance with 'opportunity to learn' as assessed by the analysis of topics from the classroom discourse.

The 'moves' within 'transactions' can indicate the types of tasks set for pupils, the guidance given and the differential treatment of pupils by teachers. They correspond to commonsense constructs of 'questions', 'directions', 'answers', and ' comments', as well as many definitions of 'teaching skills' from the research literature.

We anticipate that the results obtained from use of this instrument in the 'descriptive' and 'process-product' studies will enable us, in consultation with small conferences of teachers, to hypothesise some supplementary innovations (probably in the field of in-service training) as desirable, and to devote the final phase of this project to experimental studies of strategies to bring about these innovations.

REFERENCES

Bellack, A A; Kliebard, H.M; Hyman, R.T; and Smith, F L Jr; (1966) The Language of the Classroom; New York: Teachers College Press.

Gallagher, J J: (1970); 'A "Topic Classification System" for Classroom Observation' in AERA Monograph Series on Curriculum Evaluation No. 6 Classroom Observation; Chicago: Rand McNally and Co.

Gump, P V; (1967) The Classroom Behaviour Setting: Its Nature and Relation to Student Behaviour; US Department of Health, Education and Welfare, Final Report Project No. 2453.

Scottish Education Department, (1969) Curriculum Paper 7: Science for General Education; Edinburgh: HMSO

Sinclair, J McH.; and Coulthard, R M (1975) Towards an Analysis of Discourse; London: OUP.

ATTITUDES TO TRUANCY

ANXIETIES OVER WITHDRAWAL FROM SCHOOL: HISTORICAL COMMENT

Donald Withrington.
University of Aberdeen.

After about half a century's comparative quiescence, the subject of truancy has arisen again as a matter of concern, with working parties on juvenile delinquency setting up studies of non-attendance at school as an early warning or reflection of delinquent or feckless behaviour. Truancy has not been seen as a major anxiety to central or local government since the 19OOs and 191Os when a series of administrative changes seemed to maximise school attendance - fully free schooling, the ending of exemptions, a tightening-up of control over school age employment, the introduction of more exacting school attendance services as the units of local government increased in size and in facilities.

Thus David Rubinstein in his School Attendance in London 187O-19O4 (Hull, 1969), noting the average attendance figure of 88.2% in 19O6, goes on to say:

> The percentage attendance achieved in 19O6..has hardly been excelled in later years. Between 19O7 and the outbreak of war in 1914 the percentage in London elementary schools varied between 88.7 and 9O.1. After a drop during the war, attendance was again over 88 per cent by 1922-23. Since 1945 attendance figures have been separately calculated for primary and secondary schools in London, the former varying since 195O between 85 and 9O per cent, the latter between 85 and 92. Thus, sixty years of educational progress have been almost barren in the field of attendance.

By 19O6, there remained little concern about non-enrolments, that is, there was only a small and diminishing discrepancy between the numbers of children of school age and the numbers actually enrolled in schools - the battle on that ground had been won, after a struggle, in the course of the later 19th century. The attendance levels about which Rubinstein is commenting are the average percentage attendances of enrolled children, and include the returns of short-term absences as well as long-term ones. The London attendance levels under the school board regime, as was the case in most large and powerful school board areas, were relatively very good - better than those for many rural areas in the later 19th century (e.g. Devon or Sutherland). If we take the available figures for Scotland as a whole they show a remarkable resemblance to those for London. And this suggests that similar influences were at work in improving attendance levels throughout the country.

Averages attendance (percentaged)

Date	London	Scotland
1872	76.7	-
1877	78.7	(c.75)
1882	79.5	77.O (in 1884)
1887	78.4	77.8
1892	77.7	8O.7
1897	8O.2	83.6
19O2	83.7	84.1
19O6	88.2	87.7

After 19O6 both sets of figures are in the range 85 to 92 per cent.

What was it that had brought the improvement and stabilised the attendance levels after 19O6? And why have they not improved significantly since then?

Rubinstein sensibly sees the attainment of the high-level plateau in the returns as stemming from factors which came together in the early 2Oth century, noting that by no means all of them were directly and immediately related to schooling. Some of the worst features of poverty in the population were being removed, and the absolute need that had once existed to capitalize on the earning power of young children (for subsistence) no longer applied: in some respects this need declined as the social aid programmes (in dole and pensions, for instance) were elaborated and was lessened anyway by another major influence - the fact that social welfare services were being directly attached to the schools and could only be taken full advantage of if children were in regular attendance. Thus the distribution of clothing and boots to the poor, the very marked extension of medical services in schools, the availability of cheap or free school meals etc. were means of persuading parents to ensure the attendance of their children. (See J.M. Roxburgh The School Board in Glasgow 1873-1919 (Edinburgh 1971)). At the same time, there are signs of changes in general attitudes to children, the perceptible emergence of 'modern' teaching strategies, the building of more attractive schools, a more insistent interest in so-called 'child-centred' teaching and curricula, which seem to have added to the attraction of regular school-going: and with this some signs that employers were giving more emphasis to regular attendance and to the completion of school courses in appointing to a widening array of jobs.

In 1914, indeed, the SED annual report suggested that: "Whether there is room for further advance is open to question. An improvement in the material condition of the poorest classes and a reduction in the frequency of epidemics by sanitary precautions would doubtless have their effect. But as far as the matter depends on the willingness of the parents and the activity of the School Boards (which remained in being in Scotland until 1918), it is possible to think that the high water mark may now have been reached, and that future variations in the percentages of average attendance to average enrolments will be in the nature of temporary fluctuations due to causes beyond the control of man, such as the weather."

Yet the inspectors' reports of the day still refer to groups in the local community or factors at work which caused a higher absence from school than they thought was necessary. Some Scottish examples, from the 1912 SED report, will show what I mean. The Glasgow School Board reported an attendance of 97.9% and were smug about it. In the adjoining areas of Lanarkshire, however, we find that "attendance orders are in frequent use" and that enquiries into the causes of absence are "usually futile" among the mining and factory populations. In neighbouring Govan, the home of the Clyde shipbuilding yards, there is no comment on absence by the children of the labouring population: here it is reported that the non-attendance in the area was "probably due to the laxity of a certain type of middle-class parent who, to save trouble in his household, gives in to the whims of his children and allows casual absence from school on almost any frivolous pretext." In the Lanarkshire industrial areas, what concerned the school inspector was long-term truancy which the local authority was unable to counter in any effective way. In Goven, the district inspector worried more about the many who indulged in needless short-term, temporary, if frequent, absence rather than the comparative few who were long-term truants.

Apart from parental fecklessness, which was a common enough complaint, attention was drawn by the inspectors to the attendance problems of children whose parents moved jobs and house frequently. The 1914 report provides a suitable quotation here: "The migratory habits of certain elements of the population in the large towns constitute one of the greatest difficulties of the School Boards in these districts. The same difficulty arises to a lesser degree in the agricultural districts from the general fitting of farm servants which takes place at Martinmas or Whitsuntide over large parts of Scotland. The population of the mining districts is also highly changeable. The difficulties are accentuated in the case of tinkers and the travelling people."

But which, of all the multitude of reasons advanced in the 19th and earlier 2Oth century as explanations of truancy, are still worthy of our attention? Clearly, some have been substantially, if not entirely, removed by subsequent welfare legislation. Destitute families should not be dependant any more on the earnings of their school age children for mere subsistance, the need to pay fees or to pay for materials used in schools has gone.

The frequent appearance of epidemic disease, the liability to illness through malnutrition or insanitary home conditions or lack of clothing and shoes, these have declined or virtually disappeared. Those large, frequent and educationally-disruptive migrations of workers which so marked the 19th and early 20th centuries, whether in town or country, have largely ended. The suspicion of parents towards education because they had themselves no personal experience of a national school system (often mentioned as a factor in England) no longer applies after a century of state schooling. But there are other comments which may still be seen as applying, in some measure at least, to the incidence of truancy today: these may be summarized in three groups -

(1) Parental indifference: need of parental support.

A wide spectrum of comments was put forward here, including the declaration that working-class parents still objected to any government interference with their children, since they considered they had an 'absolute freedom to dispose of their time and labour as they saw fitting.'

It was frequently observed that legislating for compulsory schooling was not enough: for legal compulsion to be really effective there needed to be a ready alliance of state and parents who were convinced of the value to their children of the schooling that was offered (see the comment on curriculum under (2) below). Ways of gaining this parental support were hard to see or agree. From the mid-19th century onwards, however, an interesting viewpoint comes to be heard, and is exemplified in this quotation from 1857, from a speech by the principal of Homerton College: 'Unquestionably much good would result from periodical meetings of parents, at which the plans adopted for the instruction of their children might be brought under their notice and their duties and responsibilities be suitably enforced. In our early educational efforts, the relation of the parent to the school was scarcely recognized. An education deemed suitable for the humbler classes was offered for their acceptance - their views were thought unworthy of consideration, little was done to excite their interest - everything partook of a patronizing air. A very different state of things must arise ere the hearty support of our schools by the parents is secured. Our efforts must recognize the fact that none have so deep an interest in our operations as themselves, and that on their cooperation success largely depends - that we have no wish to provide education irrespective of their efforts and without reference to their wishes.... The parent brought into this intimate relation to the school and its operations would be accessible to any appeal to prolong to a certain extent the education of his child.'

The groups whom it was recognized were most difficult to reach in order to gain their support in this way were in the unskilled, frequently unemployed working-class, inhabiting urban slums noted for their persistent criminality (see comment under (3) below).

(2) Attractiveness of the school: building, curriculum, teaching methods, etc.

It was soon clear in the 187O's and 188O's that, even after tightening up legislation for compulsory schooling, absences from school were still uncomfortably high. It was one thing to get children to enrol: quite another to keep them at school. The 1851 census remarked that much parental indifference 'results from a perception of the really trifling value of a great proportion of the education offered for their parents' purchase.' Lyon Playfair gave a warning in 187O that more than a mere adding to the number of schools would be necessary: 'To compel every boy or girl to continue at school until 13 years of age involves the necessity that the instruction should aim at something more than a boy of eight may readily learn.....It would be an unredeemable hardship to compel children to attend schools unless they were made suitable to their wants in life.' The recommendation was frequently to be heard - for grammar schools as much as for elementary schools - that the content of the curricula should be decided in consideration of the occupational requirements of the locality, whether it be a seaport, agricultural area or manufacturing or commercial town.

What was taught was to be made suitable and acceptable. How it was taught was also to be carefully considered. Commentator after commentator noted that good and efficient schools in 'bad' areas, taught by sympathetic and well-prepared teachers had notably higher attendance figures than better-placed schools taught by ignorant, ill-trained or untrained, often brutal teachers. It was necessary to raise the teacher in public estimation: it was necessary to select teachers with care and to persuade them to make schooling 'delightful - attract by educative methods inside, not by compulsions outside', wrote one inspector in 1877. By implication, if not often directly, the warning is given not to alienate the children. There were pressures enough on children not to attend (for example, the taunting of a 12 year old en route to school by former classmates, now farm labourers) without adding to them in the school itself. Agricultural areas were especially badly hit by non-attendance in periods of harvesting, when child labour might be vital: many writers concluded that no amount of legal action would stem this and recommended that school holidays be so arranged to coincide with the local harvest.

(3) The separate treatment of "incorrigible" truants.

Persistent absentees often came from the lowest class for whom education often took a low priority: reasons given for withdrawing children from school were trivial. But there were truants who took days off without their parents' knowledge - or even against their direct wishes: there is a curious tale from Devon in 1886 of a boy who was strapped to the school gallery by his mother on her way to work in the morning, and a Scottish inspector in the 191O's reported that in a Moray Firth fishing town teenage boys had disobeyed their mothers, left school and gone off to sea, to earn a living and obtain their independence.

Long-term truants, the children of vagrants, orphans, street arabs and the like, all caused much heart-searching: they not only withdrew themselves from school, whenever they were got there, but they could also prompt others to leave with them. In the 186O's and afterwards in England, and earlier in Scotland, it was sometimes concluded that they should be treated separately in distinctive schools made initially more attractive to them by the availability of food and clothing: these were the day industrial feeding schools (as opposed to boarding reformatory schools, also proposed) which were similar to the 'truant schools' established in the USA in the early 185O's.

Two sets of quotations, one dated 1877 and the other 1972, may aptly conclude this paper. The first is from a lecture given to the National Association for the Promotion of Social Science by Mr. George Leith, a member of Greenock School Board, the second is from a research report by Dr. Sheila Mitchell of the University of Stirling department of Sociology (Education in the North, no. 9 (1972).

1877: The 'irregular pupil' has generally a weakly mother who requires a great deal of help. She has a number of delicate brothers and sisters who require a wonderful amount of nursing. She is very liable to catch cold, and is a martyr to 'sore heads'. 'Washing day' occurs twice in a week, and half the family is kept to assist. The irregular pupil is blessed with many relations, near and far off: and if a father's second cousin happens to visit them, the whole family is kept at home in consequence. John is absent today, because he had to go with his father's dinner, and Mary will be absent tomorrow if she has to go to the 'shop' - for two red herrings, perhaps. These are not exaggerations.

1972: We have a clear and consistent picture of the association of absence with two constellations of attitudes, on one hand the view that school is an imposition to be evaded as soon as possible, on the other hand, particularly among the boys, the pupil's self-conception of lack of academic ability. Among the girls there also seems to be an associated sense of social inferiority in some cases - being in a position where one is shamed, and scorned by schoolmates. The boys, on the other hand, are perhaps more likely to run into trouble with authority.

Poor attenders were more likely to come from families where the father was a manual worker and particularly an unskilled or semi-skilled worker. They were also more likely to come from large families. Thus the average number of siblings of the boys who were good attenders (i.e. those with 1O or fewer absences) was 2.17 whereas the poor attenders (i.e. those with 31 or more absences) had on average 3.88 siblings apiece. When we also took into account the reason for absence it became obvious that in each causal category we obtained the same picture: the children with the most

absence were those with the largest number of brothers and sisters. The largest families of all (where the respondent had an average of 4.71 siblings) were those of the boys with poor attendances whose absence was believed to be for entirely non-medical reasons.

Perhaps as a result of this larger family size the poor attenders among the boys were also found more likely to have older brothers and sisters who had left school and this was most pronounced among those who were absent frequently for strictly non-medical reasons. The presence of working older siblings in the home, presenting an example of increased personal and financial independence, could be seen as an influence away from the ethos of school towards that of the outside world.

Friendships also could be important in drawing children to, or away from, school and, here again, the poor attenders were found to spend more time with neighbourhood friends who did not attend the same school whereas the good attenders were inclined to mix outside school with classmates and other children from the same school as themselves. One 'reason' which is often suggested for school absence is that it is easier for children to play truant when mother is out at work. We could find no evidence of this since the mothers of the non-medical absentees were no more often employed than the others. We did find, however, that boys with poor attendance for purely medical reasons were more likely to have working mothers than were the better attenders - a fact which is rather difficult to interpret without further information.

SCHOOL REJECTION AND ITS AMELIORATION

John Raven.

Data relevant to the question of school rejection and its amelioration was gathered from a study involving a nationally representative sample of some 4,000 pupils most of whom were past school leaving age, and 1,200 teachers in the Republic of Ireland.[1]

These data showed that about a third of pupils sometimes or always hated going to school and felt they would be happier in a job if they left school immediately and got a job. More than half of them considered more than half of their school subjects either boring or useless. More than a quarter of the pupils thought that their school should do more to achieve 90% of the objectives we asked them about, and more than half of them felt that half the objectives deserved more attention. There would therefore appear to be a major problem of disenchantment with school.

Such levels of disenchantment make it difficult to see how teachers can achieve any goal effectively, let alone foster motivational dispositions like initiative, ability to work with others, self-confidence, and considerateness. These levels of disenchantment vary little with pupils' backgrounds, but they do vary with their anticipated destination in society. Three points deserve to be emphasised:

a. The delayed gratification hypothesis is not sustained: pupils who are going to get the best jobs like school: they are not delaying anything: they have their cake and eat it.

b. Those who are going to be teachers and administrators have no knowledge of what it feels like to be unhappy at school. They are unaware of the need for change.

c. The solution to the problem of school rejection seems to have more to do with running a series of courses suited to pupils who have different values and who will enter different sectors of society than with running enrichment programmes to adjust all to the same courses.

A detailed analysis was made of the characteristics of school rejectors. It was of very great interest - to find that, although, school rejectors had very different occupational aspirations from those who accepted school, they did not have different needs from school, wish to see different things changed in their schools, react differently to corporal punishment, or think they would respond favourably to different sorts of teachers. In our National Analysis of material collected in the course of IEA "Civics" study we found little relationship between school rejection and I.Q., but school rejectors, in addition to having different occupational aspirations to other pupils, more often than others smoked, drank, stayed out late, and got into trouble with the police.

These results, by showing that school rejection is more closely related to anticipated destination than to background, and unrelated to the other variables mentioned, seriously call into question the basic hypotheses of those who believe that school rejection is due to a clash of values between the home and the school and the basic hypotheses of those who believe that the problem is to be cured by making use of learning materials which ensure that pupils do not continue to receive failure experiences whilst at school. The intelligent do not reject school more than others. Those who intend to go into jobs for which it is irrelevant do.

Reality seems less esoteric and more down to earth than educational theorists would have us believe: If pupils do not like school it is because they cannot see the relevance of what they are doing at school to the sort of jobs they expect to enter - not because of a culture clash or because of failure experience received at school. They do not ask - and nor do many educationalists either - what schools might be doing to help them to develop the competencies they will require to perform well in the jobs they hope to enter and the life styles associated with them. As far as their perception of the objectives of education are concerned they echo the views of their fellow pupils, teachers, and parents. As we shall see later, they believe that schools should be helping them to develop their characters, but they are no more critical than others of the fact that (in the opinion of most teachers at least) schools do not happen to be doing much to achieve that objective.

Although pupils thought it was very important to pass examinations they did not think it was very important to learn about aspects of school subjects not required for examination purposes or about non-examination subjects. This contrasts markedly with their desire to have other non-examination activities on the curriculum. The inference that they not value the content of their subjects is supported by other data. They were asked "How do you feel if a teacher gives you a lesson which you enjoy but which you don't think will help you to pass an examination"? Their answers indicated that about half of them feel annoyed by such activities. Pupils were also asked if they thought there were any advantages in staying on at school after fourteen for pupils who were not taking examinations. Their replies are shown in Table 1.

Table 1

Advantages in staying on at school for non-exam pupils

	Boys %	Girls %
Many advantages	24	33
A good number of advantages	15	24
A few advantages	33	27
Hardly any or no advantages	24	12
No answer	4	3
Weighted base (= 100%)	553	573

Rather more than half of the boys, although somewhat less of the girls thought that there were few or no advantages in staying on if one was not taking examinations, providing reasonable grounds for concluding that passing examinations represents an overriding objective for many pupils.

[1] The full data is available in three volumes published by the Irish Association for Curriculum Development and the Competency Motivation Project. Further details of the results presented in this paper are available from the author.

What factors lie at the heart of the problem? Firstly, many teachers know that many of their pupils think the subjects they are teaching are boring and useless, but do nothing about it. Secondly, although, the evidence shows that there is a great deal of agreement between teachers and pupils as to what school should be doing, teachers think that examination performance, pupil guidance, vocational preparation, and the provision of other instrumentally valuable information, is a great deal less important than do pupils. They think it is much more important than do their pupils to introduce them to new academic subjects and aspects of subjects. The evidence also suggests that many teachers are more satisfied than the pupils with what they do to make their lessons enjoyable and with the amount done to encourage independence. They are dissatisfied with their pupils' achievements in the area of the 3 R's, while the pupils are relatively satisfied. Similarly, whereas we have been able to show that pupils vary considerably from one to another, and in systematic ways, in what they want out of education most teachers think that courses directed toward the same goals are equally appropriate for the more and the less academic pupils, and, whereas our data shows that some pupils are more adequately catered for than others, teachers as a rule think that the courses currently available cater equally well for more and less academic pupils. In addition, teachers markedly underestimate their pupils' serious mindedness and overestimate their concern with pop music, dancing and earning money.

However our most disturbing data shows that most teachers, while thinking that many of the objectives we have discussed were very important, do relatively little to work toward most of them and think that all but two of them are at present not very well or poorly attained.

It seems that the pupils are quite right to be demoralised: by their own admission teachers do little to work toward the goals which both they and their pupils believe to be the most important objectives of education, and they feel that these goals are poorly attained.

Why do teachers not change their courses when they know that a considerable proportion of their pupils are bored? Why do they not work toward the goals they themselves consider to be most important? What information would help them work toward these goals more effectively? What changes are required in the institutional structures in which they work?

One reason why teachers do not work toward the goals they believe to be most important is that they do not think that their priorities are shared by their fellow teachers, pupils, or their pupils' parents (Musgrove and Taylor, 1971, Kelly 1970). They live in a state of pluralistic ignorance in which they believe that they are the odd man out. If they became aware that everyone, to a very great extent, agrees with them they would be able to put their case more forcefully and be less fearful of the reactions of their colleagues, pupils, and parents to their branching out and doing something different. From the data available there is no longer any room for doubt that everyone connected with education is agreed that schools should be primarily concerned with the character development goals of education, not with conveying academic information. Yet if teachers are to work more effectively toward these goals they need to know how to achieve them, what their pupils' values and interests are, and how these vary from pupil to pupil.

The trap in which teachers find themselves is partly of their own making, and partly a result of the social placement role society thrusts upon them through the examination and evaluation system.

We have seen that teachers do not consider getting their pupils through examinations to be a very important educational goal but that, in practice, it is the goal toward the achievement of which they devote most of their energies. Although not one of their main priorities it ends up being the objective which is best attained. Furthermore a fifth of the teachers said their pupils derived no benefits from studying their subject if they did not pass the examination and, pressed to be specific about the benefits their pupils would derive, the remainder of the teachers mentioned very few, many of which turned out to be variations on the theme 'they will learn more about my subject'.

Why should it be that teachers spend their time working toward a goal they do not consider to be very important, which they believe confers few, if any, benefits, which they (inaccurately as it happens) believe their pupils do not value, and which results in many of their pupils being bored and being known to be bored?

If they recognised - as the pupils recognise, and as data reviewed by Berg (1971), Jencks (1972), and Raven (1973) suggests is objectively the case - that examination passes perform the task of allocating pupils' status positions in society, but that mastery of academic subjects does not help the pupils to lead their lives more effectively, one would have expected them to have behaved rather differently. One would have expected them to have said, "Yes, examinations are important, but, No, the study of new subjects and aspects of examination subjects is not very important".

Yet they don't do this. Could it be that they dislike, are indeed, frightened of, the idea that the social function of education is to allocate pupils' life chances with the result that, whilst refusing to discuss the implications of the social functions of education explicitly, they know in their hearts that this is what happens and therefore allow reality to influence their behaviour? And could the reason for this be that it is not only the pupils' future which is determined by examination attainments but also their own reputation as a teacher as well?

Such suggestions are supported by the data gathered. When offered the opportunity to say that examinations should evaluate progress toward the goals of education they themselves considered most important, most teachers, said that examinations should not do this. If they recognised the role which examinations - and the social selective functions associated with them - play in deflecting them from their own goals, one would have expected them to have said that they should: Making the pupils' future progress dependent on their having developed these characteristics would enable teachers and pupils to strive hard to foster them.

One reason why teachers did not think that the examination system should be concerned with assessing progress toward the character development goals of education may have been that they feared that if examinations did evaluate progress toward these goals, the teaching profession would not emerge in a very good light. There are good reasons for their suspecting this to be the case, since more than half the teachers thought that 95% of the objectives we asked them about were not very well or poorly attained: If progress toward these goals were evaluated teachers must suspect that they would be confronted with a disturbing picture of their own ineffectiveness, and that this picture would damage their public image.

Part of the explanation of teachers' unwillingness to make progress toward the goals they believe to be most important part of the selective process may be that they do not wish to recognise their own role in allocating the future life chances of their pupils. Such a task is, after all, an oppressive responsibility and a task which is not recognised when teachers are either recruited or trained. Morton-Williams et al. (1964) showed that teachers selected themselves into teaching because they wanted to do a worthwhile job in the community, to help people, and to communicate, and people with such concerns are unlikely to take kindly to the notion that their job is in reality to allocate pupils' social status and future life chances.

If these explanations of the apparent inconsistencies in teachers' responses are correct they would go a long way toward explaining why teachers find that they are unable to do much to make their subjects more interesting to the pupils. If the syllabus merely represents a set of extrinsic hurdles to be cleared by their pupils to gain entry to high status positions in society, and if both teachers and pupils are right in believing that the content of the subject is of little intrinsic value to many of the pupils, then the name of the game is to beat the system, and any attempt to enliven their teaching by bringing in additional material is to place the pupils' life chances in jeopardy. The same mechanisms could also account for the fact that teachers are unaware of their pupils' values, and underestimate their serious mindedness and responsibility. If they feel obliged to pressurise their pupils in an autocratic manner, to work toward goals which they do not themselves accept and which they believe confer few benefits on their pupils, they could not possibly trust their pupils to work toward these goals on their own. They would therefore not be able to engage their pupils in the sorts of activities which would lead to their getting to know their values and which would lead them to recognise the other qualities their pupils possess. Given the situation in which they find themselves, it would create acute cognitive dissonance if they developed a respect for pupils who possessed

qualities other than academic ability.

Finally, these suggestions go some way toward explaining the lack of pupil motivation and enthusiasm. For how could pupils be expected to be enthusiastic in a situation which teachers must themselves find soul destroying, in which neither they nor their teachers are able to work toward the goals they consider most important, and in which goals cannot be varied from pupil to pupil depending on their values and interests?

However these mechanisms do not provide a complete explanation of why teachers do not work toward the educational goals which everyone connected with education believes to be most important since two other factors are of great importance. The first is that, teachers seem to have adopted a 'single factor' model of the intellect. When asked to describe 'academic' and 'less academic' pupils, only a third of the teachers described 'less academic pupils' as 'pupils who have other qualities, pupils who are good at other things'. Two thirds of them defined them as pupils who were lazy, disruptive, disinterested, incapable of learning, and no good at anything. In other words the majority of teachers have not recognised that the qualities which came at the top of their own list of priorities are not closely related to academic ability.

The second strand relates to teachers. They define equality of educational opportunity as the same courses for everyone, and they not only think that they do provide the same courses for everyone but also that the goals are equally well attained for everyone. This finding, and its implication is disturbing for it implies that most teachers have not recognised that working toward their own goals means working toward different goals with different pupils and working in different ways to achieve the same goals with pupils who have different values and aspirations. In other words their own perceptions, both of the structure of the intellect and of the varieties of pupils, seem to represent a major barrier to their achieving their own goals. The final strand is that teachers are not acutely aware that they do not know how to work toward the goals they consider important. By international standards the IEA results show that relatively few teachers in Ireland had involved themselves in curriculum development programmes, which may indicate a lack of awareness of the problem. In addition, it is clear from our analyses that their understanding of many curriculum innovations is not very deep.

If teachers are to achieve the goals they themselves consider most important it therefore seems that they need an evaluation system which will provide recognition for both themselves and their pupils for working toward these goals, they need to be able to find out what are their pupils' values and interests, and they need to better understand how to use information on their pupils' values and aspirations to design educational programmes to foster the competencies which both they and their pupils believe to be so important.

How can the problem of school rejection be ameliorated?

Although the data points to a large number of ways in which school could be made more relevant, rewarding and enthusiasm-generating for all pupils, it shows that the lack of enthusiasm for their tasks notable in the responses of some pupils seems to be based on teachers not working toward the goals they themselves consider to be most important, on teachers not being too clear about how to go about achieving these goals, on teachers not recognising the need for diversity in types of input, and on teachers failing to take steps to deal with the conflict between the social functions performed by the educational system and their educational role. The implications for educational policy makers, lecturers in teacher training colleges, and researchers are clear. On our shoulders rests more than the burden of creating an alienated youth who will be unlikely to participate in educational programmes in the future, and squandering National resources. We have also placed teachers and pupils in an intolerable situation by not taking steps to institute more appropriate policies and courses, and by not taking steps to collect the necessary research data.

REFERENCES

Berg, Ivar (1970). Education and Jobs: The Great Training Robbery. Penguin Books, London.

Jencks, C. (1973). Inequality. Penguin Books, London.

Kelley, S. (1970). Teaching in the City. Gill and MacMillan, Dublin.

Morton-Williams, et al (1964). Undergraduates Attitudes toward School Teaching as a Career. Government Social Survey Department, London.

Musgrove, F. and Taylor, P.H. (1971). Society and the Teachers' Role. Routledge and Kegan Paul, London.

Raven, J. (1973). On Project Work and Pupil Centred Learning. Compass, Journal of the Irish Association for Curriculum Development 2 No. 1, p. 10-38.

CONCEPTUALISATION OF ATTAINMENT CRITERIA

'MASTERY LEARNING': AN ALL OR NOTHING?

Roy Sumner, N.F.E.R.

The basis for 'Mastery Learning' was first expounded by Bloom (1968) in a paper which took Carroll's (1963) 'Model of School Learning' as its conceptual paradigm. Its essence is that aptitude for learning should be viewed as a measure of the time required to carry out a learning task, rather than as defining the limits of what a student might learn. The time required would be a function of the student's characteristics and the instruction available to him. Carroll's model has been summarised by Block (1971) as follows:

	1. Time allowed	2. Perseverance
Degree of Learning = f	3. Aptitude	4. Quality of Instruction
	5. Ability to Understand Instruction	

If 1 and 3 are isolated from this expression, it will bee seen that time allowed and time required (aptitude) are the two elements that could be matched directly.

Bloom considered that the accepted normal curve of achievement was a result of uniform conditions for learning, i.e. standard pacing, expectations and instruction. If students were given 'optimal quality instruction' and the right amount of time for learning then the normal curve could be transformed into one heavily skewed towards high achievement. The strategy to bring about this state of affairs has a general framework; its implementation is what constitutes 'Mastery Learning'.

It proceeds from a realisation of what is to be learned, so that a set of 'major objectives' (i.e. content and cognitive behaviours) are defined for the students. Arising from this specification, the course is divided into a series of learning units, expected to take about two weeks of school time. Objectives for each unit are defined in relation to mastery of the major course objectives. Ordinary classroom methods are used in teaching but 'feedback/correction' procedures are employed. These are brief (often informal or ad hoc) diagnostic assessments but at the end of a unit a formative test shows what has been learned in the group situation. Supplementary instruction is then devised for those students failing to achieve the unit objectives. Formative evaluation (Airasian, 1969) is integral to teaching - learning and

informs both student and teacher of progress and difficulties. The corrective procedures do not follow a prescribed stereotyped routine but can be worked out with students to suit their own predelictions and resources to hand. Thus, small group or individual tutorial sessions might be laid on or programmed instruction at teaching machines be given.

Another schematic representation of the model is given by Anderson (1973), who points out the crucial distinction between Carroll's idea of antecedent variables (affecting learning) and Bloom's conception that these student characteristics and environmental variable can be altered. Hence:

Cognitive entry behaviours

Quality of Instruction — Time on Task — Achievement

Affective entry Characteristics

shows a chain of events. Appreciation of what students bring to a course will affect the instruction; if appropriately pitched and if poor learning is remedied the student can progress. Thus he is better fitted to enter the next learning task. Bloom, (1973) illustrates the implications (for groups) of this notion as follows:

A. Achievement distributions in which inadequate learning is not corrected at the appropriate time

Task 3

Task 2

Task 1

B. Achievement distributions in which inadequate learning is corrected at the appropriate time

Task 3

Task 2

Task 1

Figure A illustrates the accumulation of handicap suffered by the learners who were, initially, slower than the majority as well as the relative acceleration of 'high flyers'. Figure B purports to show not only that the majority of a group can achieve a high standard but that the low initial performers can surpass expectations by reaching a standard relatively close to the best in the group. Bloom (1972) makes the point that when it is known that causality is at work the educator cannot evade his responsibility for bringing about certain effects; understanding individual differences, and in particular rate of learning variations, can lead to 90% of students reaching the standard of the(previous) top 10%.

Airasian (1972) describes the outcome of a mastery learning experiment with (teacher) students on an 'introductory graduate testing methods course'. Progress tests were given after 2, 4, 6 and 9 weeks; they were graded but not used for overall assessment. Students with 75% or better were encouraged to keep to the same study pattern; the remainder were encouraged to spend more time on study or to alter their methods. The course instructor was informed of items which were incorrectly answered by half (or more) of the group (N = 35). All students were told that everyone could obtain A grades. A parallel exam. to the previous year's (essay and multiple choice) was used to compare earlier with 'mastery learning' results. Course mastery was defined as A grade (for the two years compared) and for year (1) 30% of the group obtained A's, in year (2) 80% were so classified. A study-time inventory showed that the initially poor students improved their efficiency, spending only ½ - 1 hour per week extra to achieve enhanced performance. A similar result is reported by Anderson (op cit).

Two enormously important by-products are claimed, (i) 'learning how to learn' takes place, typically with the ineffective starters improving remarkably (e.g. cutting additional time required from 5 units extra to 1), and (ii) the affective consequences of success being reinforced by success showing in better attitudes and deeper involvement. A feature of both is that competition for high grades is replaced by co-operation presumably because high grades are accessible to all.

A cynical view, perhaps, is that the arguments for mastery learning are tautological; if entry behaviours are adequate then students may learn; if corrective procedures do correct errors then students have learned; if everyone keeps up then no-one fails.

More searching criticisms are levelled by Skaalvik (1975), who identifies the strategy with programmed instruction and sees it as behaviourist in essence and therefore open to the charge that few important educational objectives - in general education- can be defined in behavioural terms. Further, he gives Atkin's (1969) argument that only a small part of the learning accruing from teaching activity can be predicted beforehand. He doubts, too, that aptitude-treatment-interaction processes are sufficiently understood for mastery learning to have its basis in cognitive, developmental stages. The strategy, therefore, is seen to be limited to a few low level tool skills as distinct from knowledge structures; and these cannot be specified as behaviours, they can only be indicated by them. Of greater moment, perhaps, is his incorporation of Skjervheim's (1965) objection to educational technology as authoritarian, rendering students the objects of manipulation. Skaalvik concludes that mastery learning ought to be restricted to teaching/learning 'basic skills and knowledge which possess hierarchical structures'.

Those of us who think along normative lines might consider that setting 'mastery levels' accessible to 80% of a student body necessarily entails accepting relatively low standards for each stage of formative evaluation. This point can be conceeded (even cheerfully) on the grounds that there is little point in setting tasks that few students can accomplish. The pedagogue's hallowed tag 'From the known to the unknown' certainly implies that the teacher should have a clear knowledge of his student's base, and in mastery learning the student also gains this knowledge (and possibly insights) explicitly.

How does mastery learning differ from 'good teaching'? Looking back at the (few) highly efficient teachers - those with their sights set on 'covering the syllabus', with each lesson organised to accomplish a visible purpose (e.g. copy out all these notes; view this film, answer these questions) - there seems little difference. They succeeded because their students/pupils as a class achieved the results expected, i.e. they all passed 'O' level, or all learned to read, or all could do arithmetic. What about the time required and corrective procedures? For those teachers the pupils were found the time and were given extra work and help. And furthermore, these teachers could teach heuristically too, but it was something of a luxury to be indulged in when the targets had been reached and the authoritarian mask could be allowed to soften. These teachers were respected but seldom liked and they demanded hard work.

Compared with mastery learning, the most significant feature is the rejection of failure, coupled with a very clear idea of what has to be done. Advanced planning and preparation is also a common feature and so is attention to sub-standard performance, indicated frequently in the marked work regularly set and punctually returned. The missing elements are that students are not given formal targets or that a variety of corrective procedures are not laid on by the teacher; pupils find their own salvation.

It would be tempting to dismiss 'mastery learning' but for the claims that the laggards improve their ability to learn. If this is a general outcome, the consequences spread far beyond mere performance in school subjects.

Mastery Learning References

Airasian, P.W. (1972) 'An application of mastery learning strategy'
Psychology in the Schools, IX, 2

Airasian, P.W. (1969) 'Formative Evaluation Instruments: A Construction and Validation of Tests to Evaluate

Learning Over Short Time Periods' abstract of unpub. Ph. D. Thesis in Chicago University, Block (op cit)

Anderson, L.W. (1973) 'Time and School Learning' unpub. Ph.D dissertation. University of Chicago.

Atkin, J.M. (1969) 'Behavioural objectives in curriculum design: A cautionary note' in Anderson R.E., Faust, G.W. and Roderick, M.C. (Eds) Current Research in Instruction, Prentice Hall.

Block, J.H. (Ed) (1971) 'Mastery Learning; Theory and Practice' Holt, Rinehart and Winston, New York.

Bloom, B.S. (1968 'Learning for Mastery' UCLA-CSEIP, Evaluation Comment, 1, No. 2.

Bloom, B.S. (1972) 'Innocence in education' School Review, 80, 3.

Bloom, B.S. (1973) 'Individual Differences in School Achievement: A Vanishing Point' Phi Delta Kuppa address to AREA Conference. New York.

Carroll, J.B. (1963) 'A Model of School Learning' Teachers College Record, 64.

Skaalvik, E.M. (1975) ' An evaluation of mastery learning' Scandinavian Journal of Educational Research, 19. 2.

Skjervheim, H. (1965) 'Eit grunnproblem i pedagogisk filosofi'. Minervas Kvartalsskrift.

THE NUFFIELD PROJECT AND ITS EVALUATION

Susan Barry.
Brunel University.

The Nuffield project was designed to meet the need created by the implementation of the Todd Committee's recommendations. Briefly the recommendations were to recognise that General Practice is a medical specialty requiring specific vocational training. The vocational training for would-be General Practitioners will consist of 2 years hospital detachment, in which the trainees work in hospital but have a 1 day detachment course in which hospital medicine is related to its general practice manifestation - and one year traineeship with a General Practitioner. In terms of manpower requirement, the implementation of this scheme means an increase in the number of General Practitioners capable to undertake one to one teaching from 300 in 1973 to 1200+ by 1977 as well as an additional number of general practitioners who would undertake the organisation and curriculum design of day detachment courses for a group of doctors in hospital training. This meant that the primary objective of the Nuffield Course would be to train course organizers, so that a mushrooming effect can be achieved throughout the country to meet the demand in such a short time.

DESIGN OF THE COURSE

The curriculum of the Nuffield Course was designed in view of the job definition of a General Practitioner as stated in the Future General Practitioner[1], since it is for the training of doctors to this job definition, that the knowledge, skills and attitudes acquired by the course organizers would be utilized.

At the inception the job of a course organizer was defined as:
"A General Practitioner who arranges courses for trainers and for trainees. He may administer a vocational training scheme. He instructs both trainers and trainees. He helps trainers to improve their management of a trainee's learning. He helps trainees to improve their management of their own learning."

With this in mind the course set up a series of Aims and Objectives. The latter defined in specific behavioural terms. Using the educational paradigm, appropriate instructional strategies to attain the objectives were designed. The teaching/learning strategies adopted were based on a design appropriate to the Doctor/patient consultation and on the general procedures of case management in general practice. Thus the Nuffield course had to offer:

1. Opportunities for heuristic learning.
2. Opportunities to distinguish process from task activities, to discuss them and to learn to use them beneficially.
3. An appropriate model of teacher behaviour and curriculum design as well as opportunities to discuss the relevance of the models offered.

It was considered that the most appropriate teaching/learning situation for the attainment of course objectives would be work in small groups with reportage and discussion at plenary sessions.

If the belief that the Nuffield Course, courses designed by course organizers, tutorials conducted between trainers and trainees, should reflect the Doctor/patient consultation(as defined in the job definition of a General Practitioner) was to be substantiated, then it seemed necessary to be able to evaluate changes which took place because of the Nuffield Course, and therefore the evaluation of the course was set up.

SHORT THEORETICAL DESCRIPTION OF THE LEARNING PROCESS

The job definition of the General Practitioner, given in the notes, makes it plain that the bulk of the objectives set up by the Nuffield project are within the affective domain and require radical changes in the attitude of the participants, who will then try and inculcate these changes to members of the courses that they are likely to organise.

At the adult level attitudes are an expression of total personality organisation. The disorganisation of attitude constellations would affect the total personality systems, for this reason the strategies adopted to produce affective changes and the resources used are of cardinal importance in the learning process.

It should be remembered that the members attending the Nuffield Course are among the elite of the general practitioner population[2], therefore it will be natural for them to attribute their eminent position and success to attitudes and behavioural patterns they held so far, thus probably creating a stronger resistance to affective change.

The Nuffield curriculum is organised in such a way that the members of the course are subjected to a severe shock to their existing attitude patterns by:-

a) putting them in the learner position.
b) creating a degree of dissonance in the affective domain.
c) creating a certain amount of anxiety which can only be resolved by trying to build new structures instead of those which have been subjected to shock.

Thus a situation is created in which a degree of both "cognitive dissonance" and "conative dissonance" is maintained. The main agent which creates and maintains the dissonance is the tutor leader of the course.

The small group within which the individual member is working provides both supportive and an anxiety creating environment. The task demands of the small group forces the exposure of certain attitudes which is anxiety creating, but by helping the individual in the completion of the task it becomes supportive. The evaluator, by being external to the set up in the knowledge field, but by being there and being internal to the learning environment, is used at this point mainly as a supportive agent.

Without discussing here theories of group dynamics and of group cohesiveness, (Stock & Thalen - Jennings - Moreno etc.) the process at the initial stages of the course is as follows. At this stage the small group coheres around hostility to the tutor leader who is the agent introducing innovation and dissonance.

Part of this hostility is neutralised by the evaluator who at that point is offered as a "sounding board" against which

dissonance can be heard and hostility neutralised. Showing some degree of neutrality in having been utilized, the evaluator, who is after all part of the external leadership, becomes a channel by which the tutor leader starts to be accepted.

As the groups fall into cohesive units within new frames of reference, they try to exclude outside resources who offer help. This has been observed to be the result of:

a) Fear of further disruption of newly accepted structures, and

b) The necessity for time to internalise and organise around new attitudes and new knowledge.

At that point the leader tutor is now identified as a peer, yet their need for approval demands that they turn to him for confirmation. When sufficient approval is not received the evaluator is appealed to for support, the evaluator is also asked to take over leadership functions and interpret between course members and the teaching resource, if the teaching resource creates some dissonance.

The dyad of tutor-evaluator work together in a way that the evaluator becomes the immediate feed-back agent and reports groups and individual anxieties, providing the leader with objective evidence against which to test his subjective judgement. In this task the evaluator parallels the role of the "sounding board" delineated vis-a-vis course members. The tutor then adjusts strategy and the degree and strength to which he can proceed in breaking structure, i.e. creating dissonance so that new learning can take place at an optimum level.

THE EVALUATION OF THE COURSE: DESIGN AND CRITERIA

As will be gathered from the short description of the learning process, the course aims at changes in very complex behavours, and a single type of research design could not measure all that was happening during the year in which the course took place.

Therefore the research design for the evaluation of the course combined several methods from formal instrumentation and measurement to that of participant observation.

The combination of research approaches and the testing of a number of hypotheses resulted in the development of a number of instruments which would measure cognitive and affective changes, which would evaluate and measure "process" in small groups and would allow the non-specific self-assessment by course members. The task of assessment also included in its design the isolation of variables which may contribute to the variance in the dimensions measured.

The following is a list of instruments used in the study:

At the starting point course participants were subjected to two tests measuring knowledge:

a) in educational theory

b) in the medical field.

3 personality tests:

a) the Eysenck personality Inventory

b) Eysenck public opinion inventory, and

c) the Myers-Briggs type indicator measuring preference on the following dimensions:

i) extraversion - intraversion

ii) sensing - intuition

iii) thinking - feeling

iv) judgement - perception

and 2 attitude tests:

a) a paired comparisons inventory, and

b) importance of general practice competencies test.

It should be noted that all these tests are dependent on criteria which has been generalized for a group of general practice teachers, i.e. they are determined by criteria external to the learner. In other words they are "philogeneric" criteria.

The knowledge tests and the attitude tests were administered once more during the course at the end of a specific learning activity and at the end of the course. In these tests a gain was measured soon after the specific learning activity had taken place. At the end of the course there was a drop on the scores in the knowledge test dealing with educational theory - yet members of the course at this time were using the educational paradigm and educational theory effectively in setting up their own courses.

In the medical knowledge test there was no drop in score but items varied in the accuracy with which they were being answered. Items which were answered correctly at the beginning were answered wrongly at the second testing and right again at the third testing etc., by the same people. 3 questions however were an exception. All course members answered these questions rightly always. Since these questions were in the medical field the question arose why. In quiry showed that the General Practitioners were very unlikely to come across the subject matter of these questions during their work and that it was based on medical knowledge acquired in medical school.

The discrepancy in the competent use of knowledge and the correct answering of formal questions on a subject raised the question of whether formal type of testing instruments were the right criteria for assessing learning and attainment.

Through participant observation, it also became painfully clear that none of the so designed instruments were really catching the flavour of the changes happening in the course so that 2 new methods of assessment dependent on criteria developed by each individual was used. In the lst method group members evaluated their peers and themselves on criteria that they considered to be important. Surprisingly there was considerable similarity in the criteria set up by individual members. These criteria as well as the ratings given by peers coincided with areas considered measurable in the tests set prior to the start of the course.

But since teaching in this course was dependent mainly on the course members' active participation in their own learning, peer assessment and self-assessment by criteria generated by the individual (autogeneric criteria), seems to be more valid.[3]

The following alteration to the job definition of a course organizer by members of the course can also be used as a self directed measure of learning. The course members modified the definition offered by course organizers which is given at the beginning of this paper to read as follows:-

"The Course Organiser is: A general practitioner who arranges courses for doctors in, or intending to enter, general practice. He may administer a vocational and/or a programme of continuing education. He instructs doctors in, or intending to enter, general practice. He helps trainers to improve their management of a trainee's learning. He helps doctors in, or intending to enter, general practice to improve their management of their own learning. He helps all doctors, himself included, to continue to improve their own learning and its management."

There is another theoretical aspect to this. All philogeneric criteria pre-suppose equality in distance gained or lost by individual members. That is to say a gain of one point or the loss of one point is measured as being of the same quality and quantity for all individuals, while the psychological distance travelled by an individual to gain that particular point might well be a multiple of the distance travelled by a different member of the group. This hypothesis was touched upon by developing a semantic differential test based mainly on Maslow's theory of self awareness. The results have not yet been analysed.

SUMMARY

The problems of assessment criteria raised by this study can be summed up as being:-

a) Can formalized test instruments be a correct measure of knowledge, after a period of intensive use, and therefore internalization of knowledge; and at what point do they cease to measure real knowledge?

b) To what extent can consensus criteria (philogeneric) accurately measure
i) individual changes? ii) process changes in group interaction?

c) How feasible and how accurate are the criteria generated by either the individual, or the group under observation, as measures of real change on which generalisations can be based?

d) To what extent the evaluation process becomes too cumbersome or expensive in combining as we have done in this study, criteria of both kinds?

The following quotations from course members' letters highlight these problems:

A Doctor who organizes a vocational course for trainees in the London area says: "It slowly comes to me how much our behaviour has changed over the last year and I am sure that changes will continue to occur over the next year or so..."

From a Doctor who organizes courses in the postgraduate medical centre in Bath: "As was said on the last day, we shan't really know what we have learned, until we have forgotten all we were taught! It will take me a while to organise in my mind the new ideas and concepts I took on board..."

From a Doctor in Dumfries who is about to become a course organizer: ".....For me the most important part of the course was what happened to me inside my head - and is still happening - rather than any immediate effect on my performance as a trainer, though that too has been affected. At times I felt a kind of Renescence Spirit.....I certainly, like Colin (another doctor in the course) had some sense of personal liberation..... "

The temptation is to go on quoting from letter after letter, which brings to our notice as evaluators of educational processes, the multiplicity of criteria by which we could measure, and the inadequacy of the criteria we have been using so confidently so far.

[1] The job definition of a General Practitioner

"The General Practitioner is a doctor who provides personal, primary and continuing medical care to individuals and families. He may attend his patients in their homes, in his consulting room or sometimes in hospital. He accepts the responsibility for making an initial decision on every problem his patient may present him, consulting with specialists, when he thinks it appropriate to do so. He will usually work in a group with other General Practitioners, from premises that are built or modified for this purpose, with the help of paramedical colleagues, adequate secretarial staff and all the equipment which is necessary. Even if he is in single-handed practice, he will work in a team and delegate when necessary. His diagnosis will be composed in physical, psychological and social terms. He will intervene educationally, preventively and therapeutically to promote his patient's health."

[2] Most members of the course were chosen by the regional advisers with the specific intention of their becoming course organizers for their region. All but 3 are members of the Royal College of General Practitioners, and in a comparative study of their medical knowledge they scored 20% more than another group of general practice teachers and their average score on an MCQ was 73%. The college examination pass mark is around 45%.

[3] A similar study was carried out among students in the Department of Surgery at the University of Miami School of Medicine by Linn et all.

REFERENCES

1. Working Party of the Royal College of General Practitioners. "The Future General Practitioner: Learning and Teaching ", RGCP-BMJ, 1972.

2. McLeish, J., Matheson, W., and Dak, J., "The psychology of the Learning Group", Hutchinson, London 1973.

3. Festinger, L., "A Theory of Cognitive Dissonance", Stanford University Press, Stanford, California, 1957.

4. Eysenck, H.J., "Dimensions of Personality", Routledge & Kegan Paul, London, 1947.

5. Eysenck, H.J., and Eysenck, S.B.G., "Personality Structure and Measurement", Routledge and Kegan Paul, 1969.

6. Maslow, A.H., "The Farther Reaches of Human Nature", Viking Press, 1971.

7. Maslow, A.H., "Motivation and Personality", Harper and Row, New York, 1970.

8. Maslow, A.H., "Towards a psychology of Being", N.J.D. Van Nostrand, Princeton, 1968.

9. Eysenck, H.J., "The Psychology of Politics". Routledge & Kegan Paul, London, 1954.

10. McLeish et al. under (2).

11. Stock, D., and Thelan, H.A., "Emotional Dynamics and Group Culture", New York University Press, N.Y., 1958.

12. Bloom, B.S., "Taxonomy of Educational Objectives", Longman, 1956 and 1964.
Handbook 1 - Cognitive Domain
Handbook 2 - Affective Domain.

13. Bloom, B.S., "The Role of Educational Sciences in Curriculum Development", Educational Sciences, Vol. I, No. I, February 1966, pp5-16.

14. Mager, R.F., "Preparing Instructional Objectives", Fearon, Paloalto, California, 1962.

15. Wiseman, S., and Pidgeon, D., "Curriculum Evaluation", NFER, Slough, 1970.

16. Flanders, "Analysing Teaching Behaviour", Addison-Wesley Publishing Co., 1970.

17. Shaw, M.E., and Wright, J.M., "Scales for Measurement of Attitudes", McGraw Hill, 1967.

18. Edwards A.L., "Techniques of Attitude Scale Construction", Ch.2, Appleton Century Crafts Inc., New York, 1957.

19. Stevens, J., "Vocational Training for General Practice: A theoretical model for an educational task analysis", An unpublished report to the Education Committee of the Royal College of General Practitioners.

20. Bernard S. Linn, Martin Arosteqv, and Robert Zeppa, "Performance rating scale for peer and self assessment", Journal of Medical Education, July 1975.

THE PROBLEM OF ATTAINMENT: PRIVATE KNOWLEDGE AND PUBLIC DISCOURSE

Stephen Kemmis.
University of East Anglia.

When we speak of a student attaining something we usually have in mind some fairly clear definition of what it is the student has attained. The concept of attainment is by its very nature object-referenced, that is, it refers (in the context of learning) to objects of knowledge. Such attainable objects of knowledge include both contents and states. For example, we may think of the attainment of a specific objective in the learning of a particular fact or piece of information (content), or we may think of the attainment of mastery of a skill or readiness for learning (state). Construing problems of learning in terms of attainment thus exerts a powerful influence on the way one can think about learning: it focusses attention on the content or subject-matter to be learned, and thus emphasises learning outcomes defined by reference to such content. In general, the notion of attainment leads us to think of problems of teaching and learning in the realm of public discourse, that is, as problems of learning subject-matter which can be explicated or referred to in language.

Much of the psychology of learning has dealt with "learning" defined in these public discourse terms. In particular, behaviouristic psychology has preferred to characterise learning in terms of publically observable student performance. By doing so, behaviouristic psychology has located the criteria of learning and thus of attainment in the realm of public discourse.

Three key features of the rise of instructional psychology from behaviouristic psychology predispose it towards a public discourse characterisation of the criteria of learning and attainment:

(1) its methodological preoccupation with publically observable behaviour,

(2) its view of knowledge as transmissible content or information and

(3) its view of education as instruction which has led to its preoccupation with the mastery of content.

These features of instructional psychology lead to a conflation of private knowledge and public discourse and, in doing so, impose severe limitations on the usefulness of instructional psychology to educators.

EDUCATION AS DIALECTICS

Education is a dialectical process in which the student moves between private knowing and public discourse. When Peters and others speak of education as "initiation" of students into a "public world" they seem to recognise the dialectical process by which private knowing is moulded under the influence of public discourse.

If psychology is to be of use to educators, it must address this dialectic. To do so, it must reject the three key features of instructional psychology identified earlier.

First, it must pay more careful attention to the knowing subject who participates in the dialectic of education, rather than defining the criteria of attainment solely in terms of what is to be learned in the realm of public discourse. Instructional psychology recognises the importance of learner characteristics in the language of entry behaviours, but defines these by reference to the terminal behaviours to be attained after instruction which are, in turn, characterised by reference to subject matter-based performance criteria. It also recognises the importance of changes in the learner in its attention to learning as relatively enduring changes in performance and transfer. But instructional psychology literally cannot conceive of the learner or of changes in the learner independent of the learning tasks: its language is limited to characterisations of the learner in public discourse terms. To redress this imbalance it is necessary to attempt to describe the private knowledge of the learner.

Second, a more nearly "educational" psychology must revise its view of knowledge. The objectivist view of knowledge as transmissible information suggests that learning is a process of building representations of subject-matter in the head of the learner. Such a view fails to account for the dynamic quality of thought. Furthermore, it suggests a crude correspondence theory of knowledge and truth: if these representations are some sort of copy of the external world (or public discourse), then how could we ever decide on their adequacy as representations? In order to decide on their adequacy we would need some homunculus-like being or process which has access to the world independently of the copy. The workings of this homunculus would then be the thing to be explained.

Private knowledge may be conceived as the cognitive structures of an individual. The development of knowledge is a constructive, dialectical process of interaction between the cognitive structures of the individual and the world (including public discourse). Cognitive structures are manifest in action upon the world (i.e. interaction with it); they are characterised descriptively, by inference from the formedness of action. Cognitive structures are not inert, guiding performance by remote control; they are dynamic and unfold in action. In this view, knowledge inheres in action and is revealed by it. It does not precede action. (Action is not mere performance, it includes intention and thought).

Third, a more nearly "educational" psychology must consider the dialectic between the knowing subject and the world of public discourse with which he is confronted as part of the process of education. Rather than attempting to bring the student's performance into conformity with criteria of attainment of public knowledge (for example those underlying a criterion referenced test), the educator will create a sympathetic resonance between the forms of private knowledge of the student and the forms of public discourse under discussion. The educational psychologist may study the nature of this resonance as revealed in the interaction of student and subject matter. If sympathetic resonance is to be achieved it is not sufficient to study public discourse formulations of learning tasks, although it is necessary to do so. It is also necessary to understand the forms of the student's knowledge which are by their very nature obscured from public vision and enormously various. Learning is always individualised in the sense that the student interacts in his own characteristic ways with the subject matter; individualised instruction has attempted to match student performances on entering instruction with appropriate subject-matter, but has taken the form of the public discourse to be unproblematic (e.g. fixed by the "structure of the discipline" or a logically-derived learning hierarchy), and the nature of student deviation from that form to be essentially a matter of preparation in content terms. But the student must be prepared in terms of cognitive structure (i.e. must have forms of action which can assimilate the new subject matter) not just content.

By taking into account the knowing subject, by viewing knowledge as inherent in action and revealed by it, and by viewing education as a dialectical process in which the cognitive structures of the student are brought into sympathetic resonance with public discourse, we can take a broader view of the problem of attainment. It is, at the very least, more complex than instructional psychology allows. An "educational" psychology must break out of the constraints imposed upon instructional psychology by its history. An examination of more complex forms of action may reveal the necessity for doing so; an examination of the methodology of structural analysis can reveal the possibility of doing so.

STRUCTURAL ANALYSIS: A NEW APPROACH

As an example of a complex form of action, we may take clinical decision-making. The teaching of clinical decision-making has often relied on statistical decision theory (originally explored extensively in economics). Although the model of the decision situation it assumes is most often used to describe the behaviour of decision-makers or to prescribe more rational procedures for decision-making, the model is inadequate in explaining or even characterising the decision-making process in psychological terms; that is, it is outcome, not process-oriented. Moreover, it depends upon the elements (outcomes, probability, utility) of the decision process, the alternative actions available, and relevant information all being explicit. If these features of the process

are explicit, and if the aim of the teaching is to shape the student's behaviour so that it more closely approximates the rational process embodied in the model, then an instructional theory founded upon operant conditioning would be sufficient in teaching clinical decision-making.

The clinical decision-making situation departs from this ideal model in two ways, however:

(1) the decision-situation is rarely explicit: the elements of the decision process depend upon the tacit knowledge of the diagnostician, the alternatives are generated by the decision-process (they do not precede it except in a logical sense), and the relevant information is highly ambiguous and

(2) the process of reaching the diagnosis is not usually "rational", it proceeds by intuition and "best guessing". The critical feature of clinical decision-making as regards diagnosis is strategy; the capable diagnostician must develop structures of action which are highly sensitive to the emerging structure of the case at hand. The traditional teaching of diagnosis by apprenticeship suggests that the process is largely tacit; specialisation in medicine suggests that different diagnostic structures underlie different physiological and disease systems.

To teach clinical decision-making, therefore, a more complex model seems to be required. Such a model must allow for the existence of tacit knowing and intuitive strategy. Although analyses of the clinical decision-making process are not currently available, they would seem appropriate. An educational psychology of the kind outlined above may provide the basis for teaching clinical decision-making.

The problem of teaching such complex subject-matters as clinical decision-making may yield to better understanding of the nature of the activity itself. In less esoteric fields, structural analysis has begun to yield insights into cognitive structure which have immediate implications for education. Several examples of research into cognitive structures involved in education (being carried out at the University of Illinois) suggest the promise of the method:

(1) in a traditional Piagetian conservation situation, Kamara (1971) has identified several types of cognitive structures which underpin the general conservation structure:

(a) operational schemas like the scheme for reversible actions

(b) object manipulation schemas like those for "flattening" and "balling" plasticene, and

(c) perceptual structures to do with identifications of shapes and comparisons of surface areas;

(2) Easley (1972) reports work by Triplett which identified several structures utilised by a child in incorporating adult conceptions of heat into his own conceptions; the structures have to do with the child's notions of expansion, heating and the movement of molecules,

(3) Leiser (1974) has identified a number of different notions of division utilised by twelve-year-olds, grouped together by the equivalence of their result, they are (a) long division, (b) "invert and multiply" - the algorithm for fractions, (c) divide A cakes among B persons, (d) how many times does a slice of size B fit into A cakes, (e) move the decimal point to divide by powers of ten, (f) group theory, and (g) a list of facts without any deep meaning: e.g. 34 divided by 2 equals 17;

(4) Witz and Easley (1972), on the basis of work carried out by Rosalind Driver, were able to develop the notion of "physical deep structures" which may be important in understanding a range of adult and child behaviour. These structures are particularly obvious as intuitive kinesthetic feelings aroused when we think, for example, of the concept of force, and reach out into our understanding of force as a physical and social concept.

These examples suggest that the way students think about subject-matter may be vastly different from the way teachers conceive it, but also that analysis can reveal pedagogically-relevant structures in their thinking. Furthermore, the identification of these structures in the action and language of children can suggest ways in which the teacher can engage and develop children's conceptions by posing new problems.

On the basis of their work in structural analysis, Witz, Goodwin and Easley (1974) have suggested a cognitive model for the evaluation of units of instruction which is compatible with the dialectical view of education presented here.

While the concept of attainment is object referenced, the problem of attainment for educators and educational psychologists is more than that. To some, the alternative appears to be a flight to subjectivity - an alternative which raises more serious problems than the view it replaces. The present paper has advocated an interactionist position which attempts to understand the process of interaction between the cognitive structures of the student and the reality which confronts him. In education, much of this reality is the public discourse of the subject-matter. The problem of attainment can be conceived in terms of "sympathetic resonance". But to avoid the danger of replacing a problem with a mystery, it is necessary to find ways of identifying this resonance. Structural analysis provides such a method by developing accounts of the cognitive structures which are revealed in the child's interaction with teachers, other students or objects. In general, human action is remarkable for its smoothness and coherence, as various kinds of objects are assimilated into existing cognitive structures, but these structures may not be compatible with adult conceptions: the problem of identifying attainment is a problem of locating incompatibilities between private knowledge and public discourse, and posing the problems which allow the child to develop more encompassing structures.

[1] This paper is a version of a longer paper on cognitive structure and decision-making being prepared under the auspices of the UNCAL (Understanding Computer Assisted Learning) evaluation of the National Development Programme in Computer Assisted Learning.

REFERENCES

Easley, J.A. "The Structural Paradigm in Protocol Analysis", Journal of Research in Science Teaching, 1974, II, 281-290.

Kamara, A.I. "Cognitive Development among School Age Themne Children of Sierra Leone", Ph. D. dissertation, University of Illionois, Urbana, 1971.

Leiser, D. "Semantic Content and Arithmetical Systems - Two New Cognitive Structures Identified", unpublished Master's Thesis, University of Illionois, Urbana, 1974.

Witz, K.G, Goodwin D R and Easley, J.A. "A Cognitive Model for the Evaluation of Units of Instruction" Instructional Science, 1974, 3, 307-326.

Witz and Easley, J.A. "Cognitive Deep Structure and Science Teaching", in K G Witz and J A Easley, Analysis of Cognitive Behaviour in Children, 1972, Final Report, Grant No. OEC-0-70-2142(508), USDHEW, NCERD.

A BRIEF SURVEY OF THE SCOTTISH EXAMINATION SYSTEM

EXAMINATIONS IN SCOTLAND: 1960–75

J. D. Wilson.
Moray House College of Education.

The Scottish Education Department (SED) introduced external examinations for secondary schools in 1888. Since 1965 responsiblity for their administration, and for advising the Secretary of State for Scotland on related matters has lain with the Scottish Certificate of Education Examination Board (SCEEB). It currently supervises two examinations at the Scottish Certificate of Education - ordinary grade (1962) and Higher grade (1888) - as well as the Certificate of Sixth Year Studies (1968). On the advice of the Board the Dunning Committee was set up to review arrangements for assessment at 16+, when 'O' grade is normally taken by school candidates, and to consider the implications of any changes for subsequent examinations. At the same time, on the recommendation of the Consultative Committee on the Curriculum, the body which at least in theory is charged with oversight over the whole school curriculum, the Secretary of State has set up the Munn Committee to consider the structure of the curriculum in S3 and S4, the years which lead up to presentation at 'O' grade.

From the above, it is clear that centralised control is a feature of Scottish examinations, that the curriculum and the examination system are seen as two separate entities, and, perhaps most significantly, that there is an awareness that all is not well with present arrangements.

STATISTICAL DATA.

The number of pupils staying on at school and taking certificate examinations has increased rapidly in recent years. In 1966 there were 56, 414 pupils in S4 - S6; by 1974 there were 120, 272. The table below shows the trend to more candidates, more presentations and more examinations:-

Year	Number of Candidates	Number of Subject Presentations	Number of Papers Lower/Ord.	High er.	CSYS
1888	972	3, 000	6	6	-
1961	18, 500	88, 700	20	23	-
1962	42, 300*	218, 200	37	35	-
1974	125, 168*[1]	544, 442	45	42	16

* External candidates at further education centres were admitted to S.C.E. exams for the first time in 1962.
1 S.C.E. candidates only.

The sharp rise in presentations between 1961 and 1962 reflects the impact of 'O' grade. The percentage of school leavers gaining any S.C.E. qualification has gone up from 43.4 per cent in 1970-71 to 57.7 per cent in 1973-74. In 1974 78.8 per cent of pupils in S4 were presented for examinations at the Ordinary grade.

THE EXAMINATIONS

a) S.C.E. 'O' Grade

Concerned at the fact that only one in three pupils admitted to S.L.C. courses was securing a certificate an Advisory Committee in 1959 recommended the introduction of a fourth year examination whose standard was such that 'a pupil who is at the lower end of the top 30 per cent of any age group should, with satisfactory teaching and adequate effort on his part, have a reasonable prospect of securing passes on the 'O' grade in at least three subjects in the fourth year". (SED 1959). Initially a pass/fail examination, 'O' grade results have been awarded on five bands since 1973, with bands D and E corresponding to marks of 40-49, and 30-39 per cent respectively.

b) S.C.E. 'H' Grade

Established in 1888 the 'H' grade is recognised as the senior school award of standing for university matriculation requirements. Passes are awarded on three bands. Pupils achieving in the mark range 40-49 per cent may be awarded a compensatory pass at 'O' grade.

c) Certificate of Sixth Year Studies

CSYS was introduced to provide a goal for a 'proper sixth year' for able candidates who had obtained the necessary S.C.E. qualifications for matriculation. Candidates may be presented in up to three subjects in which they already hold 'H' grade passes, and are awarded a certificate in which each subject is graded on an A to E scale. The certificate does not count as an entrance qualification for higher education, and unlike S.C.E. courses the curricula for CSYS are not referred for approval to the Scottish Universities Council on Entrance. The aims of CSYS are set out in the Examination Boards Annual Booklet as "to assist the schools in promoting study in depth and the capacity for the independent pursuit of a subject, in advancing the educational maturity of those who already have attained a recognisable proficiency in a subject discipline and in giving direction and focus to study undertaken by pupils in the sixth year of Scottish secondary schools".

THE S.C.E.E.B. AND ITS WORK

SCEEB has a monopoly in examining at the secondary level. The Board is representative of a wide range of interests - schools, colleges, universities and industry - affected by its examinations. The important work of syllabus construction, setting and conducting examinations is devolved to small subject panels and examining teams which include teachers, HMI and College and University staff. The subject panels 'ensure that developments in schools arising from teacher initiative, or through the influence of various curriculum development centres are suitably reflected in the Board's syllabuses and examinations' (SCEEB, 1971). The idea of partnership is strengthened, too, by the involvement of schools in the conduct of examinations; principal teachers (subject) invariably have a major say in the presentation of candidates on the basis of results in one (sometimes two) internal preliminary examinations, and teachers' estimates of performance are forwarded to the Board along with the names of candidates. The Board goes to great lengths to ensure that examinations are seen to be conducted fairly. In 1974, 11, 468 appeals were considered at S.C.E. and over one third were allowed.

The precise nature of the Board's relationship to the S.E.D. and the Consultative Committee on the Curriculum has intrigued many students of the Scottish educational scene. In Scotland, as in England, the examination system is an instrument to define syllabuses, establish standards and motivate and influence the teaching and learning strategies of teachers and pupils. Macintosh (1970) noted that 'by writing the examination syllabus we avoid writing the teaching syllabus'. In its report on Secondary Education the Scottish Advisory Council outlined the effect an external examination could have on the structure and quality of education: "When standards are still insecure and a supply of teachers adequate in scholarship and in professional training has to be built up slowly over one or two generations, a wisely directed external examination may be a potent instrument for good. Its defined syllabuses, its proper balance of emphasis, the consistency of its level - all these exert a steadying influence and a salutary stimulus on the teachers". (S.E.D. 1947). Both 'O' grade and CSYS have had a similar impact. It is no exaggeration to say that the former has been an important factor in changing the structure and organisation of secondary education, and SCEEB syllabuses for 'O' and 'H' grade have transformed the curriculum by splitting up broad fields of knowledge, such as science, into separate subjects.

McPherson and Neave (1974) have argued that a main reason

for the introduction of CSYS was to change teaching and learning strategies in sixth year by causing the introduction of projects and laboratory work in science and tutorial work and dissertations in arts subjects. The fact that both 'O' grade and CSYS could be introduced only three years after they had been initially proposed has led many, especially South of the Border, where the saga of sixth form reform drags on, to admire the Scottish pattern of speedy, and apparently uniform, implementation. Scrutiny of the proposals of the CSYS subject panels shows, however, marked differences in the way the general aims of the certificate were translated into practice, and there is evidence that some CSYS courses are taught formally with teachers using dictated notes and handouts, and exercising close supervision over their pupils. It is curious, too, that the objective of promoting independent study should have been so publicised for CSYS, with the implication that this was not an appropriate objective for SCE candidates.

While the day-to-day work of the Board has been widely praised many aspects of its operations may legitimately be criticised. The Board has been reticent in promoting discussion of the functions of examinations in Scottish education. It does not appear to have considered whether it has a role in validating school courses developed for non-academic pupils, and several schools have gone to C.S.E. examining boards for guidance which SCEEB might have supplied. More seriously, perhaps, it has done little to educate teachers, or the public, in the limitations of its awards. A system of externally set and conducted examinations provides few opportunities for teachers to become sophisticated in examination techniques. Its syllabuses - with the exception of Physics - are set out in a way which suggests unfamiliarity with the literature on curriculum and the definition of objectives. The impressionistic reports of its chief examiners are vague and subjective, and ultimately quite unhelpful. But it is the lack of statistical data which is the most serious weakness of the Board's operations. It has been acknowledged (Currie, 1972) that 'the SCEEB are a major branch of Scottish educational research. They are obliged to make a continuous assessment of the relevance and validity of their own assessment procedures'. The statistical data produced annually by the Board requires close study, and since even percentages are not always given, it is not clear, without doing the sums oneself, how different groups of candidates compare. No evidence is published on the standard error of the examinations, and without such evidence it is impossible to know whether reliability is constant from year to year.

The Board's prime concern has been to demonstrate that standards are being maintained at a time of rapid growth of presentations amongst school and further education candidates: its motto significantly is "In trutina ponentur eadem". In the report for 1971 the Board's duty in this respect was noted, but it was "recognised that it was not, and never had been the policy of the Board in fulfilling this duty or judging its degree of success to do so simply by reference to the proportion of candidates obtaining passes. Changes in education and social conditions, and in parental attitudes were in themselves likely to increase the number of candidates with the required potential". (SCEEB, 1971). In the event the pass rate at 'O' and 'H' grade has plummeted recently as the figures below indicate, and school candidates are now achieving passes at the level of candidates in further education.

'O' Grade	School Candidates	All Candidates
1969	76.0	73.8
1974 (A-C Grades)	61.9	61.9

'H' Grade	School Candidates	All Candidates
1969	75.2	72.6
1974	65.1	64.0

Percentage pass rate at 'O' and 'H' Grade S.C.E. in 1969 and 1974.

But there are other considerations in defining and maintaining standards which the Board does not seem to have taken into account. Does the Board, for example, lay down common criteria for determining pass rates or for awarding different bands of pass in different subjects? What account is taken of the ability of candidates taking a particular subject? Weir (1975a) and Kelly (1975) suggest that more able pupils are presented for subjects like latin, physics and german (at both 'O' and 'H' grades) while less able candidates take technical and domestic science. At the same time it appears that the more able the candidates the more severely the subject is marked. A Schools Council (1974) study has come to similar conclusions. Weir (1975b) found that 49 per cent of 'O' grade latin candidates gained band A passes in the 1973 examination, but so too did 45 per cent of the intellectually heterogeneous group who sat arithmetic. Ought the difficulty of the subject to be taken into account? How does one decide whether physics is more difficult than woodwork? To what baseline should marks be standardised if grading differences are sufficiently large - to the severe or the lenient standard?

In assessing the work of the Board credit must be given for its role in curriculum development and for the innovations in examining it has promoted. Multiple choice tests and oral examinations are now regularly employed. But, perhaps inevitably in a central body, it has not sought to create that scepticism of the reliability of its awards which would be considered appropriate by an unbiased observer who looks at the effect of the examinations on school practice and the attitudes and behaviour of teachers, pupils and others who set store by the awards.

THE FUTURE OF 'O' GRADE.

The separate existence of SCEEB reinforces the impression that in Scotland the examination system and the curriculum are separate entities. It is important to stress that assessment is an element in curriculum, and that decisions about assessment techniques and standards of attainment can only sensibly be taken in the context of discussions about the objectives of education and the content and methods through which these objectives will be realised. In this paper discussion of the objectives of education in S3 and S4 would be out of place, but as age 16+ is increasingly regarded as the end of general education a pertinent question is the structure of course and standard to be aimed at.

Looking back it is clear that in the last fifteen years we have seen the definition of a new standard of general education to be aimed at by the majority of secondary pupils. Until 1951 award of a group leaving certificate, of four or five higher or lower passes, on a course approved by S.E.D., was normally equated with successful completion of the secondary stage. Such an award was gained by less than ten percent of an age group, and it is debatable whether the standard of general education should be pegged to university matriculation requirements. The introduction of 'O' grade brought an attainable and worthwhile goal closer for a much larger proportion of the school population.

Of course 'O' grade was a departure from traditional Scottish practice and necessitated a restructuring of the secondary course with selection of subjects being made at the end of S2 rather than S3. Able pupils were still expected to embark on a three year course for 'H' grade presentation in S5, though they might take subsidiary or additional 'O' grade subjects in fourth, fifth or sixth year. (In practice the majority took all their 'H' subjects at 'O' grade in S4). Other certificate pupils began a two year course leading to 'O' grade in S4: for them it was a terminal objective, although an increasing number were persuaded by their success to stay on to S5 and S6 to attempt 'H' grades. Pressures from non-certificate pupils quickly built up and with comprehensive education, ROSLA and the introduction of banding the exam now attempts to provide an objective for the whole ability range. 'O' grade has also served as a popular initial qualification for students at further education colleges, although it may be questioned whether it is a suitable examination for mature candidates.

It would appear that Scotland has created a common system of examining by accident while researchers in England have suggested that it is almost impossible to do so by design. The figures quoted earlier show, however, that the examination takes a heavy toll of candidates. Less than 2 in 3 presentations are successful, and 38 per cent of 'O' grade attempts by S4 pupils in 1974 resulted in failure. Part of this fail rate can be explained by 'hopeless' presentations of pupils, some of whom lack the ability even to read the question

paper, but pupils presented in a large number of subjects also contribute to the total, and this may reflect the increasing divergence of 'O' and 'H' grade syllabuses in some subjects, or wholesale presentation of very weak candidates in the hope that they will get D and E grades and scrape a B or C grade in the easier subjects.

There is widespread agreement that the nature of assessment at 16+ needs to be rethought. S3 and S4 are increasingly seen as the second cycle of general education which should be conceived as a unit for all pupils. A variety of courses to suit the varying abilities of pupils needs to be recognised. Schools at present put pupils into certificate courses, knowing that they will be lucky to get an E band because they believe employers will give preference to pupils who have followed such a course. Against this kind of competition there is little incentive for teachers to devise suitable non-certificate courses. Significantly schools which have developed such courses are increasingly going for C.S.E. validation to obtain evidence of national currency for their pupils.

After S4 pupils are faced with a range of options involving continued study at school, or in further education, or entry to the world of work. The function of assessment at such a stage is not primarily selective and allocative, but rather to further the self knowledge of the pupil, and to guide him and provide evidence for teachers, parents and employers so that the best decisions can be taken. Some award of national standing for all pupils, perhaps banded to reflect levels of performance, on a variety of two year courses internally devised, but with an element of external moderation would gain wide support. It might be suitably supplemented by a profile of the pupil's skills, interests and non-intellectual attributes, provided such ratings could be made sufficiently reliable. The Headmasters' Association of Scotland is currently piloting such an instrument. Such an assessment is remarkably similar to that recommended by the Advisory Council in 1947.

CERTIFICATES AFTER 16+.

The aims of the post 16+ curriculum and their assessment, are topics for separate papers. Here the discussion is confined to the present functioning of 'H' grade and CSYS courses, and draws heavily on the work of McPherson and Neave (1974) who have surveyed qualified school leavers in Scotland in 1970 and 1972.

The traditional strength of 'H' grade has lain in the fact that it was tied to university matriculation requirements and these guaranteed a broad secondary course. Both claims would be challenged today. Although most pupils take mathematics and a foreign language at 'H' grade, choices in S2 have a profound effect on the university faculty the pupil enters. Even pupils who follow a course with a genuine arts/science mix are much more likely to enter science than arts faculties. CSYS confirms this specialisation: very few CSYS candidates study subjects across the arts/science divide, even if they have taken mixed courses at 'H' grade.

In our inflationary world university entrance requirements have risen steadily, partly out of a need to select applicants, but partly too as a guarantee against allegedly falling 'H' grade standards. Entrance standards have risen highest in arts and social science faculties where the competition for a place is keenest. In a situation where more pupils are acquiring the minimum qualifications and the available number of places is not increased, the "going rate" can only get higher.

Alternatively additional academic qualifications are sought. It is possible that this will be the fate of CSYS.

A long-standing aspiration of Scottish education was to provide a 'proper' sixth year of advanced study for pupils who had secured university entrance qualifications in fifth year. CSYS was designed to provide an objective for such studies. In practice the proportion of pupils taking CSYS courses by themselves has been small: only 1 in 10 of all sixth year leavers, and 1 in 5 of university entrants. The majority of candidates were taking new subjects at 'O' or 'H' grade or repeating subjects to secure a better grade for university entrance purposes. Given the continuing pressure on places it is possible that certain university departments will treat CSYS as an additional entrance qualification, which will solve their dilemma temporarily but in the long term merely add to the hurdles pupils must negotiate to reach university. Unfortunately there is little evidence that CSYS improves the level of first year performance at university. Perhaps this indicates merely that, for most CSYS students, the impact of genuinely innovatory and independent learning strategies is weakened by the continued exposure to traditional pedagogy in S.C.E. courses. Alternatively it may be that universities do not reward the student who has been trained to work by himself.

CONCLUSION

In the last fifteen years social and economic pressures have acted on the Scottish education system to create a much bigger 'pool of ability'. Structural barriers within schools have been lowered and the system has moved significantly towards a 'contest' model (Turner, 1960) where all pupils have access to certificate courses, and further opportunities to qualify are available for people in employment.

We are moving rapidly towards the credentialled society, but the credentials seem to be worth less than before. Minimum entrance qualifications have become meaningless, and it is the 'going rate' which counts for university entrance. In employment, on the other hand, the well-qualified now compete for jobs which demand only two or three 'O' grades. Yet school qualifications are disappointing predictors of subsequent performance at university, and what does possession of three 'O' grades - 2 C's and a D - really tell an employer about an applicant?

What are the purposes of assessment? We seem to be pre-occupied with the use of examination results. Selection is a problem for the selector: we should educate him to understand the limitations of the results we provide. Grades can only tell so much: we should encourage him to ask for other data about our pupils. We tend to use examinations to motivate the pupil and determine the syllabus and forget that their main function is to measure achievement and improve teaching. Perhaps we ought to think more in terms of criterion-referenced testing: examinations should tell us (and the pupil) how far we have obtained our objectives. Effective teaching and assessing cannot be separated. In theory at least teachers are responsible for the curriculum, but they will never exercise that responsibility effectively until they have become its assessors too. Internal examinations, properly conducted, can be as potent a force in improving the quality of an education system as external examinations: SCEEB, in association with the universities, colleges and the curriculum development centres, has a role to play in developing the necessary expertise among teachers.

References

Currie, N.B. (1972) Several Shades of Grey TESS 15th December, 1972.

Kelly, A. (1975) Standards in Highers Paper presented to SRHE (Scotland) May 10th, 1975.

Macintosh, H.G. (1970) The Place of Tests in Secondary Education in Ingekamp (1970), Developments in Educational Testing, Vol. 2 ULP.

McPherson, A.F. and Neave, G (1974) An Improper Sixth Year Centre for Educational Sociology, University of Edinburgh

S.C.E.E.B. (1971) The Scottish Certificate of Education Examination Board and its Work. Gibson

Schools Council (1974) Examinations Bulletin 29 Comparability of Standards between Subjects Evans/Methuen Educational.

S.E.D. (1947) Report of the Advisory Council on Secondary Education, HMSO.

S.E.D. (1959) Report of the Working Party on the Curriculum of the Senior Secondary School, HMSO

Turner, R. (1960) Modes of Social Ascent through Education: Sponsored and Contest Mobility, Am Soc. Rev. 25(5)

Weir, A.D. (1975b) Bending the O Grade TESS 1st August, 1975.

Weir, A.D. (1975a) The Scottish Certificate of Education: Factors Affecting Pupil Performance Scottish Educational Studies 7 (1) 5-14.

THE RELATIVE STANDARDS OF SUBJECT EXAMINATIONS

Alison Kelly.
University of Edinburgh.

This paper describes the results of some comparisons of the relative standard of examinations in different subjects. The comparisons involve Scottish O and H grades over the seven year period 1969-1975. The results are highly consistent from year to year to year, and between O and H grades. The patterning of results is similar to that found by other researchers in a study of English GCE and CSE subjects. A strong hierarchy of subjects exists, in terms of both the standard of the examinations, and the average ability of the candidates attempting each examination. Languages are at the top of this hierarchy, followed by sciences, social subjects and vocational subjects. There is a marked tendency for the more difficult examinations to be attempted by the more able candidates. When the results are analysed separately for girls and boys, sex differences are evident in the difficulty levels. Girls find languages easier than do boys, but boys find sciences easier than do girls.

INTRODUCTION

There is widespread intuitive belief that standards between and within subjects do vary, but a little thought shows that the question is far from simple. What is meant by standards? In what sense can they be compared? Is it relevant or useful to compare different subjects? Is standardization a practical proposition? The Schools Council and several English examination boards have sponsored some research on these questions over the past few years, but until now there have been no large scale studies giving a time series of results.

If entrance requirements for further education or employment were always specified in terms of subjects, it would perhaps not matter if some subjects were, in some sense, easier than others. But in fact requirements are often expressed as 'four O grades' or '2 B's and a C', implying that passes in any subject are equally acceptable, and represent a similar achievement. Indeed a fundamental equivalence between subjects is implied by the generic names 'O grade' or 'H grade'. So, on the grounds of fairness to pupils with qualifications in different subjects, who may be in competition for the same position, some effort should be made to establish a real equivalence.

Standards in different subjects can only be compared if they are defined in a way which is not specific to any particular subject. For the purposes of comparison it is useless to say, for example, that candidates in Mathematics must be able to differentiate a simple expression, and candidates in French to use the subjunctive, since differentiation and the subjunctive are fundamentally incomparable. A definition which can be applied to any subject has been spelt out for the CSE Grade 4: 'a 16 year old pupil of average ability who has applied himself to a course of study regarded by teachers of the subject as appropriate to his age, ability and aptitude may reasonably expect to secure Grade 4'. Other examinations may not be aimed at 'a 16 year old pupil of average ability' but similar, norm-referenced definitions could be devised and used as a basis for standardization.

If all candidates took all subjects and were marked by the same examiner, standardization could be achieved by equating the distribution of marks in each subject (assuming the candidates were equally motivated and industrious in all subjects). Since this situation does not exist, statistical procedures can be used to approximate it. The aim of such a standardization is to ensure that no candidate is penalised by the marking style of the examiners in a particular subject, or by the calibre of the other candidates taking that subject. Clearly the candidates taking some subjects are, on average, more able than the candidates taking other subjects. A standardization procedure will compensate for this by taking the calibre of candidates into account when setting the distribution of marks in any particular subject.

DATA, METHOD AND ASSUMPTIONS.

The work reported here was carried out for the Scottish Certificate of Education Examination Board, using their records of results in Ordinary ('O') and Higher ('H') grade examinations for the seven years from 1969 to 1975[1]. In addition the results of Highers examinations taken by fifth form candidates in 1962 were available from the Scottish Education Data Archive (Armstrong & McPherson, 1975). Complete records were used, so no sampling was necessary. In these records candidates' marks are recorded on a fourteen point scale, which was reversed for the purposes of analysis so that a high score corresponded to a high mark. The reversed scale of banded marks ran from 14 (over 90% in raw marks) to 1 (less than 30%) in bands of 5% of raw marks.

The method of analysis has been described in detail elsewhere (Kelly, 1975[2]) and will only be outlined here. In essence the procedure is to compare the mean banded mark candidates obtained in any one subject with the mean banded mark obtained by those same candidates in all other subjects. The difference between these means is taken as a measure of the standard of the subject in question. The mean banded mark in all subjects serves to establish a consensus standard against which the standard of a particular subject can be compared. For example if candidates who attempt French do worse, on average, in French than in other subjects, then French is considered a difficult subject. (The terms 'easy' and 'difficult' are used throughout this paper as a convenient shorthand to describe standards. A more exact formulation is 'leniently graded relative to the consensus standard' and 'severely graded relative to the consensus standard'). However the mean grade candidates obtain in other subjects is dependent upon the standard of those other subjects - if the other subjects are easy, candidates will obtain better mean grades than if the other subjects are difficult. There is considerable evidence that candidates tend to take groups of easy or difficult subjects together, and to compensate for this tendency an iterative procedure is used. The first difference between means (described above) is added as a correction to all marks in the subject, and the calculation repeated using these corrected banded marks. This is repeated until all new differences are less than a specified size. The total correction then indicates the standard of the examination in that subject, relative to a consensus standard.[3]

The main assumption underlying this method of standardization is that there is some form of general academic ability, so that candidates who do well in all subjects will, on average, also do well in any particular subject. The corrected mean banded mark in all subjects is used as an estimate of this general ability. For example, candidates who take Latin tend to gain above average marks in all other subjects; their average mark in Latin is therefore adjusted to a high level, on the assumption that these are able candidates and that if all candidates (including the less able) attempted Latin the mean mark would fall. Factor analysis and intercorrelations suggest that the general ability assumption is valid for academic subjects, but less so for artistic and vocational subjects. This assumption and others upon which the method is based are discussed more fully in my earlier paper (Kelly, 1975).

RESULTS

(a) Time Series

Tables 1 and 2 show the corrections for the principal O and H grade subjects over the last seven years. Only O grades taken from fourth year and H grades taken from fifth year have been considered. In the Tables positive corrections indicate difficult (severely graded) subjects, whereas negative corrections indicate easy (leniently graded) subjects. The size of the correction specifies the amount by which banded marks in that subject would need to be adjusted to bring the subject into line with the consensus standard. Unit correction is equivalent to one band, or 5% in raw marks.

The most striking feature of these tables is the stability of the corrections from year to year. With only a few exceptions, subjects which were difficult in one year tended to be difficult in other years as well, and to differ from the consensus standard by approximately the same amount. For H grades, where results are also available for 1962, this basic consistency has existed over a period of fourteen years. Where there have been changes in relative standard these have, in general, been

TABLE I. Corrections for 'O' grade subjects, taken from SIV I975 - I969.

Subject	1975	1974	1973	1972	1971	1970	1969
English	-0.68	-0.53	-0.64	-1.09	-1.40	-0.65	-0.61
History	0.78	0.53	0.13	0.11	0.43	0.49	0.60
Geography	0.84	1.10	0.47	0.47	0.28	0.63	0.55
Modern Studies	-0.27	-0.61	-0.34	0.40	0.20	0.31	0.40
Arithmetic	-0.45	-0.35	-1.07	-0.57	-0.74	-0.88	-0.80
Mathematics	1.46	1.17	0.79	0.29	0.64	0.45	0.13
Latin	0.56	0.45	1.22	0.61	0.65 /	2.13	1.73
Greek	0.67	1.18	1.47	1.50	1.72 /	2.94	2.93
French	1.56	1.22	1.34	1.44	1.01	0.39	0.48
German	1.96	1.88	1.35	1.65	1.46	1.13	0.97
Italian	1.99	1.54	1.57	1.42	0.76	0.47	0.33
Russian	2.25	2.09	1.48	2.59	2.47	2.79	1.68
Spanish	1.69	1.29	1.31	1.17	1.30	1.08	0.56
Gaelic (Learners)	-1.28	-1.47	-0.84	-1.37	-1.17	-0.16	0.32
Art	-0.87	-0.96	-0.29	0.08	-0.04	0.07	-0.08
Music	0.22	0.07	-0.29	-0.09	-0.43	-0.32	-0.35
Economics	0.30	0.56 /	0.74	1.08	1.37	1.13	0.63
Secretarial(Shorthand)	-0.28	-0.33	-0.22	-1.02	-0.44	-0.78	-1.34
Secretarial(Audio)	-0.88	-1.05	-1.07	-1.82	-1.65	-1.78	-1.66
H.E.(Food & Nutrition)	-1.89	-1.80	-1.55	-1.10 /	-1.92	-1.58	-1.36
H.E.(Fabric & Fashion)	-0.82	-0.22	-1.00	-0.89 /	-0.85	-1.32	-1.08
Anatomy, Physiology & Health	-0.10	0.08	1.41	1.66	0.77	0.85	0.92
Building Drawing	0.18	0.12	1.31	0.47	0.35	0.60	1.88
Woodwork	-3.39	-2.73	-2.21	-1.76	-1.54	-1.56	-1.74
Metalwork	-3.62	-3.80	-2.44	-2.21	-1.90	-1.74	-2.51
Engineering Drawing	-0.55	0.06	.08	0.11 /	0.69	0.33	0.19
Biology	0.55	0.51	0.63	0.24	0.37 /	-0.03	0.07
Chemistry	0.46	0.28	0.65	0.77	0.51	0.45	0.70
Physics	0.47	0.09	0.59	0.35	0.84	0.19	0.77

NOTES:
I) Positive corrections indicate difficult (severely graded) subjects; negative corrections indicate easy (leniently graded) subjects. Unit correction corresponds to one band ie 5 o/o of raw marks.
2) Corrections for I972 to I969 have been scaled by a factor of I.33 to compensate for a change in the standard deviation of banded marks. This scaling does not affect the rank order of subjects, but allows corrections before and after the change of scale to be compared.
3) An oblique stroke (/) between two columns indicates a change of syllabus between those years.

TABLE 2. Corrections for 'H' grade subjects, taken from SIV - I975 - I969 and I962.

Subject	1975	1974	1973	1972	1971	1970	1969	1962
English	-0.33	-0.37	-0.35	-0.40	-0.49	-0.52	-0.46	-0.11
History	-0.32	-0.32	-0.06	-0.03	0.14	0.31	0.31	0.67
Geography	-0.41	-0.42	-0.36	-0.25	-0.28	-0.18	-0.24	-0.06
Modern Studies	0.05	-0.12	0.09	0.50	0.65	0.24	0.41	-
Mathematics	0.81	0.68	0.56	0.38	0.51	0.13	-0.15 /	+0.36
Latin	-0.64	-0.07	0.51	0.92 /	1.40	1.24	1.17	0.91
Greek	-0.45	0.96	0.41	1.13 /	2.24	2.19	1.70	1.01
French	0.55	0.59	0.40	0.48	0.29	0.27	0.28	0.62
German	1.42	1.41	0.99	0.85	0.68	0.68	0.58	1.53
Italian	0.56	0.44	0.60	0.59	0.65	0.42	0.13	1.03
Russian	1.02	1.41	1.34	1.21	1.35	1.45	0.96	0.94
Spanish	0.63	0.80	0.82	0.73	0.73	0.34	0.24	0.83
Gaelic(Learners)	-1.42	-0.91	-0.75	-1.86	-2.48	-1.70	-1.44	-0.85
Art	-1.07	-0.51	-0.58	-0.33	-0.31	-0.35	-0.59	-0.23
Music	-0.68	-0.64	-0.56	-0.65	-0.25	-0.71	-0142	-0.92
Economics	0.43	0.55	0.95	0.85	0.87	0.71	0.42	-
Secretarial (Shorthand)	-0.72	-0.59	-0.66	-0.90	-0.73	-0.05	-0.68	-
Secretarial (Audio)	-1.21	-1.10	-0.89	-0.82	-0.88	-0.56	0.07	-
H.E.(Food & Nutrition)	-1.24	-1.11	-1.74 /	-1.42	-1.65	-1.24	-0.93	-1.14
H.E.(Fabric & Fashion)	-0.43	-0.53	-0.57 /	-0.89	-0.78	-1.07	-1.09	-2.00
Anatomy, Physiology & Health	0.73	0.77	1.21	1.19	1.16	0.88	0.07	-
Building Drawing	-1.06	-1.21	-0.71	-0.76	-0.49	-0.61	0.19	-2.71
Woodwork	-2.57	-2.21	-1.99	-2.02	-2.17	-1.47	-1.69	-
Metalwork	-2.83	-2.09	-1.88	-2.34	-1.84	-1.58	-2.02	-1.28
Engineering Drawing	-1.41	-1.02	-1.10	-1.12	-1.22	-1.37	-1.11	-1.34
Engineering Science	0.74	0.73	0.76	0.30	0.72	1.18	0.89	-
Biology	0.71	0.78	0.70	0.88 /	0.80	0.97	0.70	-
Chemistry	0.19	0.20	0.29	0.33	0.44	0.49	0.66	-
Physics	0.21	0.08	0.02	0.04	0.13	0.25	0.45	-

NOTES:
I) Positive corrections indicate difficult (severely graded) subjects; negative corrections indicate easy (leniently graded) subjects. Unit correction corresponds to one band ie 5 o/o of raw marks.
2) An oblique stroke (/) between two columns indicates a change of syllabus between those years.

TABLE 3. Comparison of the difficulty of each subject and the ability of the candidates taking that subject at O and H grade in I975.

Subject	'O' grade		'H' grade	
	Difficulty (correction)	Ability (corrected mean banded mark)	Difficulty (correction)	Ability (corrected mean banded mark)
English	-0.68	5.88	-0.33	6.15
History	0.78	7.39	-0.32	6.41
Geography	0.84	6.99	-0.41	6.22
Modern Studies	-0.27	5.09	0.05	5.30
Arithmetic	-0.45	6.36	-	-
Mathematics	1.46	8.04	0.81	7.04
Latin	0.56	10.22	-0.64	7.90
Greek	0.67	10.27	-0.45	8.63
French	1.56	8.43	0.55	7.24
German	1.96	9.27	1.42	7.81
Italian	1.99	8.18	0.56	7.39
Russian	2.25	9.76	1.02	7.77
Spanish	1.69	7.91	0.63	7.18
Gaelic	-1.28	6.45	-1.42	5.33
Art	-0.87	5.20	-1.07	5.24
Music	0.22	7.88	-0.68	6.51
Economics	0.30	6.58	0.43	5.99
Secretarial (Shorthand)	-0.28	5.84	-0.72	5.20
Secretarial (Audio)	-0.88	4.40	-1.21	4.56
H.E.(Food & Nutrition)	-1.89	3.97	-1.24	4.58
H.E.(Fabric & Fashion)	-0.82	4.53	-0.43	4.88
Anatomy, Physiology & Health	-0.10	4.54	0.73	5.35
Building drawing	0.18	4.95	-1.06	4.80
Woodwork	-3.39	3.60	-2.57	3.91
Metalwork	-3.62	3.79	-2.83	4.00
Engineering Drawing	-0.55	5.55	-1.41	5.23
Engineering Science	-	-	0.74	6.85
Biology	0.55	6.57	0.71	6.36
Chemistry	0.46	8.04	0.19	6.98
Physics	0.47	8.08	0.21	6.83

NOTES:
I) The difficulty of each subject is measured by the corrections given in tables I and 2. The ability of the candidates taking each subject is measured by the corrected mean banded mark those candidates obtained in all the subjects they attempted.

TABLE 5. Corrections for O and H grade subjects in I975 calculated for boys and girls separately, with the difference between the corrections for boys and girls.

Subject	'O' grade			'H' grade		
	Correction		Difference	Correction		Difference
	Boys	Girls		Boys	Girls	
English	0.08	-1.39	1.47	-0.04	-0.59	0.55
History	0.64	0.89	-0.25	-0.51	-0.21	-0.30
Geography	0.49	1.37	-0.88	-0.44	-0.35	-0.09
Modern Studies	-0.51	-0.08	-0.43	-0.18	0.24	-0.42
Arithmetic	-0.85	-0.07	-0.78	-	-	-
Mathematics	1.17	1.18	-0.71	0.66	1.11	-0.55
Latin	1.10	0.25	0.84	-0.23	-0.84	0.61
Greek	0.82	0.64	0.18	-0.24	-0.61	0.37
French	2.37	0.96	1.61	0.98	0.37	0.51
German	2.72	1.69	1.03	1.43	1.36	0.12
Italian	2.02	1.99	0.03	0.95	0.40	0.55
Russian	3.37	1.63	1.74	2.01	0.69	1.32
Spanish	2.33	1.32	1.01	1.08	0.40	0.68
Gaelic(Learners)	-0.02	-2.39	2.37	-	-	-
Art	-1.04	-0.75	-0.29	-0.92	-1.22	-0.30
Music	-0.48	0.46	-0.94	-1.43	-0.46	-0.97
Economics	-0.01	0.80	-0.81	0.38	0.60	-0.22
Secretarial(Audio)	0.41	-0.99	1.40	-	-	-
H.E.(Food & Nutrition)	-0.51	-2.06	1.55	-	-	-
Anatomy, Physiology & Health	-0.52	-0.09	-0.43	0.36	0.70	-0.34
Engineering Drawing	-0.57	0.50	-1.07	-	-	-
Biology	0.11	0.78	-0.67	0.46	0.82	-0.36
Chemistry	0.32	0.81	-0.49	0.07	0.53	-0.46
Physics	0.28	1.36	-1.08	0.03	1.11	-1.08

NOTES:
I) The corrections indicate the difficulty of each subject.
2) In the 'Difference' columns positive scores indicate subjects which were relatively easier for girls than for boys; negative scores indicate subjects which were relatively easier for boys than for girls (NB This does not give any information on the actual level of achievement of boys and girls in the subject).
3) Only subjects taken by at least twenty candidates of each sex are included in this table

FIGURE I. The relationship between the difficulty of O grade subjects and the ability of the candidates attempting each subject in 1975.

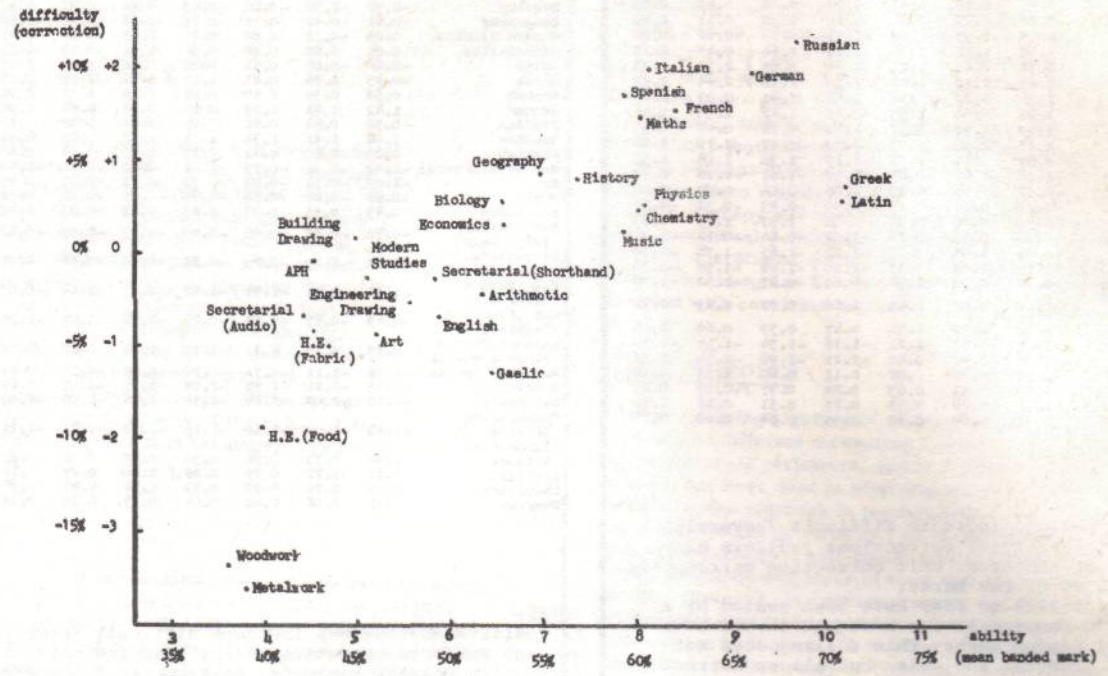
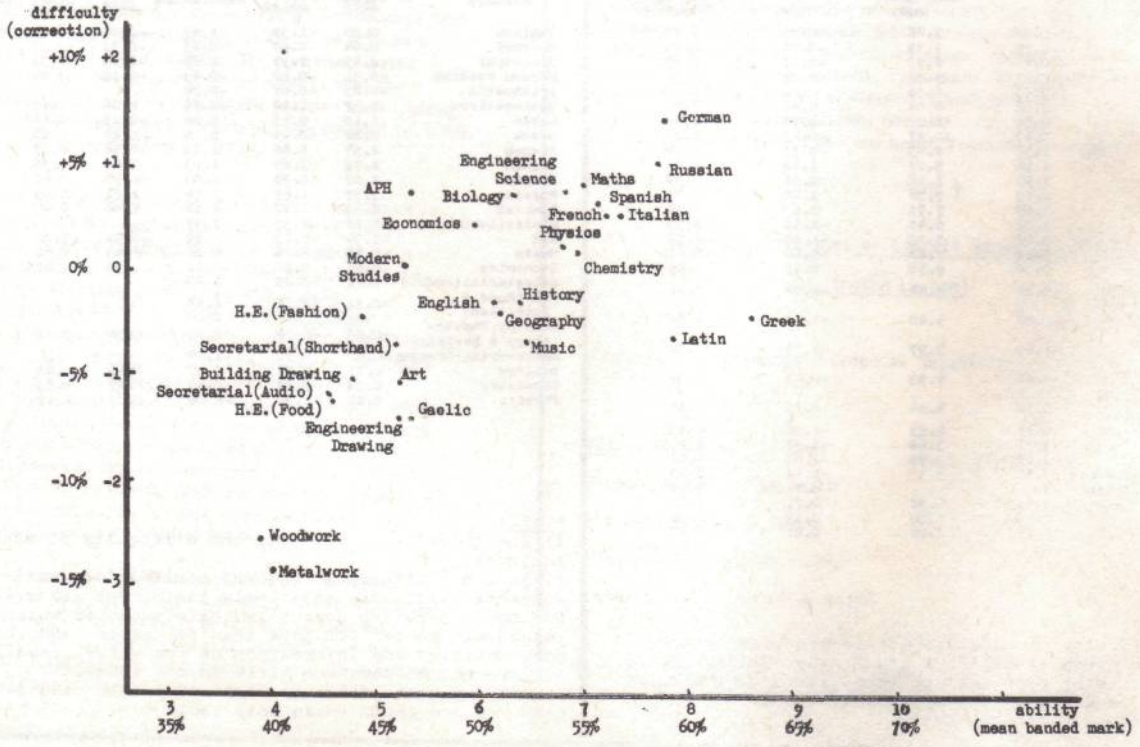


FIGURE 2. The relationship between the difficulty of H grade subjects and the ability of the candidates attempting each subject in 1975.



gradual changes rather than sharp discontinuities. This is well illustrated by the steady trends over the past seven years in O grade Greek (which has got relatively easier) and O grade French (which has got relatively more difficult).* Perhaps the most startling changes have been in H grade Latin and Greek, which have altered over the past two years from being the most difficult subjects to being easier than the consensus standard.

* It must be emphasised that these corrections only give the relative standard of different examination within any one year. They say nothing about any changes in absolute standard.

(b) Hierarchy of subjects

Two different hierarchies of subjects can be established: a hierarchy of difficulty, and a hierarchy of ability. These systems are shown in Table 3 and Figures 1 and 2. The hierarchy of difficulty is obtained from the corrections, and measures the difficulty of each subject relative to the consensus standard. The hierarchy of ability is obtained from the corrected mean banded mark of the candidates attempting each subject, and measures the average general ability of those candidates.

Even a casual glance at Figures 1 and 2 shows that these two hierarchies were closely related. By and large the most difficult subjects were modern languages, followed by sciences and social subjects, with vocational and practical subjects much easier. Until recently classical languages have been even more difficult than modern languages. There was a similar ordering of ability. The most able group of candidates were those who took classical languages, followed by modern languages, sciences, social subjects and vocational subjects. Thus subjects taken by more able candidates were, on average, more difficult than subjects taken by less able candidates. At both O and H grade the Pearson product-moment correlation between the difficulty of a subject, and the average ability of the candidates taking that subject was large, ranging over the seven years between O.58 and O.72 at O grade, and between O.61 and O.84 at H. grade. The most noticeable anomalies were Gaelic at O grade which, over the last five years, has been consistently easier than might be expected from the calibre of its candidates, and Anatomy, Physiology and Health at H grade which has been consistently more difficult than might be expected. In 1975 Latin and Greek were highly anomalous at both grades, being relatively easy, although attempted by extremely able candidates.

Both hierarchies are widely spread. In 1975 the standard of O grade subjects ranged from 11% more difficult than the consensus (Russian) to 18% easier than the consensus (Metalwork) - a spread of nearly 30% in raw marks. The range at H grade was slightly less (from +7% to -14%) but still considerable. And the examinations cater for candidates over a wide range of general ability. Those attempting Greek at O grade in 1975 averaged a standardized mark of 71% in all their examinations, compared to those attempting Woodwork who averaged a standardized mark of only 38% in all examinations. At H grade the range between the same two subjects was from an average standardized mark of 63% to an average standardized mark of 40%.

Stability is the hallmark of these results. As already noted, the corrections are similar from year to year. This means that the hierarchy of difficulty is maintained from year to year. The same is true of the hierarchy of ability in which the yearly changes have been even smaller (not shown). There is also a close resemblance between the hierarchies at O and H grade. Pearson product moment correlations between the corrections at O and H grade for the past seven years are around O.85, while the correlations between the average ability of the candidates taking each subject at O and H grade are over O.95.

But the stabilities extend even more widely. Table 4 compares the average rank order of difficulty of nine subjects at O and H grade with the average rank order of difficulty found by Nuttall, Backhouse and Willmott (1974) in a similar investigation of English GCE and CSE grades in 1968. In spite of the differences in date, level and educational system, there is still a basic similarity between the ordering of subjects. Art and English are near the bottom in all cases, and, with a few exceptions, French, Mathematics and the sciences are near the top. Although not included in the summary tables, Nuttall's results also showed that practical and vocational subjects were leniently graded and that other languages (German and Latin) were severely graded. A strong positive correlation between the difficulty of a subject and the ability of the candidates attempting that subject was also apparent in the English examinations and there was a similar hierarchy of ability between the subjects.

TABLE 4: The average rank order of difficulty of nine subjects in Scottish O and H grade examinations and English GCE and CSE examinations

	Scottish		English	
Subject	O grade	H grade	GCE	CSE
Art	8	9	8.5	9
Biology	7	1	5	5
Chemistry	4	4	1	3
English	9	8	8.5	8
French	1	3	2	1
Geography	3	7	7	6
History	6	6	4	7
Mathematics	2	2	6	2
Physics	5	5	5	4

NOTES

(1) The Scottish results are the rank orders for these nine subjects derived from Tables 1 and 2, and averaged over the seven years 1969-1975.

(2) The English results are derived from Nuttall et al (1974), Table VIII 3, pg. 63 by averaging the rank orders for the four boards within each sector.

(c) Sex differences

When the examination marks are analysed for boys and girls separately, the corrections obtained are strikingly different. Table 5 shows this for 1975, and the results for other years are similar. The difference between the correction for boys and for girls shows which subjects are easier for each sex, relative to an assumed common zero. It is noticeable that girls find languages, Home Economics and Secretarial Studies easier than do boys; but boys find science, mathematics and social subjects easier than do girls. This is not to say that girls actually do better than boys at languages, or boys better than girls at sciences; only that, relative to their general ability, they do better. In fact the boys who do languages and the girls who do science tend to be, overall, more able than their counterparts of the opposite sex and to do equally well (not shown). These patterns have changed very little over the past seven years, and again Nuttall et al (1974) found similar patterns for sex differences in England.

DISCUSSION

These results suggest several points which will be mentioned rather than discussed here. The relative standards of subjects have stayed fairly constant over a period of seven years, and this suggests successful standardization within any one subject from year to year. But the relationship between difficulty and ability raises questions about the relationship between an examination and its constituency. Do able pupils attempt difficult examinations because the less able could not cope? Or are the examinations made more difficult in order to stretch the pupils? Perhaps neither, since anomolies persist. H grade Anatomy, Physiology and Health has been consistently more difficult than might be expected on the basis of the ability of the candidates attempting it, and yet neither has the examination got relatively easier nor have the candidates attempting it become relatively more able. If Latin and Greek continue as relatively easy examinations it will be interesting to see whether this produces any change in the average ability of the candidates attempting classics. Possibly the constituency is determined more by the future destination of the candidates - APH for nursing, classics for university - than by the difficulty of the examinations.

Criteria-referencing retains some attractions despite arguments for norm-referencing. Perhaps English appears easy and foreign languages difficult because the candidates have been learning English for so much longer than foreign

languages? This argument stresses attainment rather than potential. It is more compelling in the case of different forms of the same syllabus - to take a hypothetical example, the same mathematics syllabus could be taught using workcards and using conventional methods. The examinations would take a slightly different form, but it would be easy to see whether candidates had reached the same standard in mathematics. But supposing candidates actually learnt more mathematics by conventional methods (or conversely by workcards). Standardization of these examinations would result in a pass being awarded for more learning (although equal potential) on one syllabus than another. Fair or not?

This type of speculation could continue. Are some subjects intrinsically more difficult than others - are their concepts more difficult - irrespective of the standard of the examination? Are examiners tempted to increase or decrease the popularity of their subject by adjusting the standard of the examination? But to finish with a more concrete problem: what causes the sex differences in difficulty level? By H grade at least it is unlikely to be a question of motivation. Those boys who are studying languages and those girls who are studying sciences have chosen these subjects for specialist study; they are highly self-selected; and yet they find the subjects more difficult than their counterparts (less highly selected in terms of numbers) of the opposite sex. Why?

Acknowledgments

This research was funded by the Scottish Education Department and the Scottish Certificate of Education Examination Board. However, the views expressed are those of the author alone.

1. Scottish 'O' grades are usually taken for the first time at the end of fourth year of secondary school, and are generally considered to be similar in standard to English 'O' level. s 'H' grades are usually taken for the first time one year later, and are considered to be intermediate in standard between English 'O' and 'A' levels.

2. Copies available from the author.

3. For ease of computation with large numbers of candidates the iteration was in fact replaced with an exactly equivalent linear equation solution (described in Kelly, 1975).

4. It must be emphasised that these corrections only give the relative standard of different examination within any one year. They say nothing about any changes in absolute standard.

References

Armstrong Christine & McPherson Andrew (1975), 'SEDA Dictionary of Variables for SCE H grade qualified school leavers in 1962, 1970 and 1972'

Kelly Alison (1975), 'A Study of the Comparability of External Examination in Different Subjects'

Nuttall D.L., Backhouse J.K. & Willmott A.S. (1974), 'Comparability of Standards Between Subjects' Schools Council Examinations Bulletin 29, Evans/Methuen Educational.

TEACHER PARTICIPATION IN EDUCATIONAL RESEARCH

ENLISTING THE PARTICIPATION OF TEACHERS IN RESEARCH

Daphne Johnson.
Brunel University.

Teachers have always been called upon to help in the carrying out of educational research, usually by completing questionnaires or by administering tests or questionnaires to their pupils. Research projects of this kind do not involve teachers and researchers in a working relationship over an extended period of time - indeed, if the distribution of the research material is undertaken by the L.E.A., teacher and researcher may never meet (1). Whilst impersonal studies in the quantitative genre remain a valuable source of data, the vogue for participation has its influence in educational as well as in other research spheres. If the research design requires the active participation of teachers in the analysis and reporting of findings, the researcher needs to take account of hierarchies and working relationships in the school. This paper is chiefly concerned with the influence of such organisational features on teachers' participation in research. Before turning to these factors, however, some aspects of research design based on the clinical model will be discussed.

The Clinical Model of Research

In contrasting the 'clinical' with the 'engineering' model of research design, Gouldner (2) has stressed that the clinician's basic commitment 'is not to a particular therapeutic technique, but rather to a distinctive role definition'. In Gouldner's terms, the consulting 'engineer' conceives and completes his research assignment largely in terms formulated by the client whereas the 'clinician' recognises that he has the problem of helping his client learn something.

Writing in 1965, Gouldner anticipated the recent proliferation of action research, much of which invokes the clinical model, but he expressed some reservations about the practical possibilities of securing acceptance of the clinical model in relations with groups, as distinct from individuals. Nevertheless, he commented that from work then proceeding at the Tavistock Institute it appeared that these difficulties would not prove insurmountable. The research procedure of social analysis used by the Brunel Institute of Organisational and Social Studies builds on the Tavistock experience, and the recently formed Educational Studies Unit is pioneering the use of a variant of this research procedure in school-based research. Collaborative analysis involves participants (for the purposes of this discussion, teachers) in individual discussions with researchers about those aspects of their role which have relevance for the particular area of research interest. Researchers prepare written role reports which they discuss and amend with individual participants. The reports are then cleared for circulation to other participants in the project who occupy contingent roles. In a series of joint discussions, participants and researchers work towards the preparation of a composite document which will draw out common themes and give expression to any differing perceptions of the researched situation. Such documents are used as information for policy-making either within the institution or outside it.

It will be seen that the procedure of collaborative analysis relies on the active and continuing participation of members of the institution under study, and requires them to interact not only with researchers but with each other. In school-based research, the researcher or research team who follow this procedure will be working in the school over an extended period. With the establishment of large comprehensive schools, sub-systems of considerable complexity have increasingly been required to co-exist within a single institution. Educational researchers have, as yet, paid little attention to the implications of this organisational complexity, either for the achievement of educational objectives (3), or for the viability of all types of school-based research. But when there are almost one hundred teachers on the staff, some will be closely linked in working relationships, while others are

almost unknown to one another. The researcher in a school of this size quickly becomes aware of the need to acquaint himself with and be sensitive to the priorities of sub-groups within the staff community.

If the research enquiry focusses specifically on aspects of school organisation (4), the participation of teachers involved in many of the school's sub-systems is likely to be required. So far as collaborative analysis is concerned, the validity of the emerging data rests on the extent to which individual subjective accounts can be objectivised by setting them alongside one another in a discussion climate which permits open reference to their variation. Useful triangulation is achieved if the perceptions of senior, middle management and junior staff can be recruited (5).

The participative mode of action research is not, of course, confined to organisational enquiry (6), nor to the particular research procedure of collaborative analysis (7). Any research team who hope to enlist the participation of teachers must give due weight to the importance of negotiating processes, and take account of the involvement of members of staff in hierarchies and working relationships which may influence or be affected by their participation in the research.

L.E.A. Approval

It has, of course, always been the case that any research approach to a particular school must be preceded by negotiations at l.e.a. level. This is no less the case for research on the clinical model. As Kogan and van der Eyken (8) have indicated, Chief Education Officers back off from any suggestion that they 'manage'head teachers and their schools, each institution being allowed to develop an organisational and educational style of its own. But so far as entry into a school by an outside group such as a research team is concerned, the prior approval of the CEO is vital. In the case of a research programme which is likely to have implications for policy-making outside the institution under study, some mechanism for on-going legitimisation of the research is desirable. For example, a Steering Committee, chaired by a senior educational administrator, may be given the task of sanctioning the research design and successive stages of the research programme. Such a Committee will greatly strengthen the credibility of the research team in their negotiations within and between institutions.

The Head Teacher as Gate Keeper to the School.

Although the CEO's prior approval is essential, the research team will be aware that head teachers stand at the gate-way to their particular school. Neither the CEO nor the governors can require a head teacher to admit researchers to his (9) school, so when an empirical study seeks the positive collaboration of teaching staff the key nature of the role of the head teacher cannot be overemphasised. The subsequent contribution of all other potential research participants within the school will depend on the head teacher's positive response to the research team's request for access to the school. In negotiating such access, researchers must take account of factors which may motivate or inhibit agreement by the head teacher. An obvious positive motivation might be the head teacher's confidence and pride in his school and especially in the particular aspect of school life to which the research relates. But the head teacher may still be prepared to allow research access even if he is dissatisfied with the particular aspect of school life on which the research focuses, if the research design follows the clinical mode. The head teacher is more likely to be willing to uncover possible weak areas of school performance if he is being asked to collaborate in a participative research programme rather than if he is being asked to provide data for a descriptive study of the school, in the subsequent interpretation of which he will play no further part. The head teacher will also be positively motivated to allow researchers access to the school if he judges that the research is likely to be of a quality which will enhance the school's position in some way, by giving prestige, by improving staff or pupil performance (the 'Hawthorne' effect so often observed as a spin-off from research activity), or by attracting local authority interest and possibly resources.

Factors which are likely to inhibit the head teacher from allowing researchers access to the school are, of course, equally varied. Setting aside the pathology of personality (for whilst paranoid head teachers are to be found in our own, as in any other, educational system, they are fortunately an exception) the research negotiators may find themselves up against a general disquiet on the part of the head teacher about the extent to which the school is meeting its objectives, with the result that 'outsiders' seem an unnecessary complication of an already difficult situation. Or there may be apprehension about the probable reaction of sub-groups within the school (I wouldn't mind, but they would'). Whether or not this is a displacement of the head teacher's own anxieties will only become apparent if it proves possible to open negotiations with the groups in question. The judgement that the research will not prove to be of a quality which will enhance the school's position in any way is an inhibiting factor perhaps more likely to be encountered by the lone researcher than the research team. On the other hand the head teacher may form the judgement that a high powered or prestigious research team are likely to prove insensitive to the task-priorities of teachers or pupils. Hopefully, most of these inhibiting perceptions can be brought out and talked through during the negotiating process which is so important a preliminary to field work.

Research Participation by Teaching Staff

Possible motivations for the head teacher to allow researchers access to the school have been very generally described, but since active participation in a lengthy research process cannot be conscripted, the motivation of individual teachers must also be considered.

If the head teacher has become fully convinced of the desirability of the research programme, he may well bring formal or informal pressure to bear on other staff to participate. From the point of view of subsequent relationships between the research team and participants, such pressure is undesirable, but in any case teachers will need some personal motivation if participation is to continue over time. Often such motivation will be provided by a concern about the particular aspect of school life to which the research project relates - a concern which is not necessarily based on knowledge, but perhaps on an uneasy awareness of ignorance. When the philosophy of research in the clinical mode is explained, teachers quickly appreciate that high level expertise in the area under research is not a pre-requisite of participation, since the whole research process will provide a learning situation for all concerned (10). If it is envisaged that the findings will feed back into professional or in-service training, many teachers may find this an additional motivation to participation, especially if they have become conscious of gaps in their own preparation for the teaching task.

The tendency in many schools to play down the role of accountability and control within the institution may also be a factor in eliciting teacher participation in research, especially the participation of junior staff. In the experience of the writer, the young teacher is much less inhibited about contributing his individual and newly-minted point of view to the research study, than a similarly junior member of a more overtly bureaucratic structure is likely to be. The extent to which young teachers become aware of the lines of authority and accountability to which they are subject will, of course, vary from school to school, and even between departments within a single school. Where the head of department actively assumes a managerial role vis-a-vis the staff in his department, junior teachers will readily acknowledge their accountability to him, and only through him to the head teacher. Where the department is kept in close touch with the specialist Adviser, some sense of responsibility to the latter may also be felt and expressed. But in many cases it appears that teachers leave colleges or institutes of education with no conception of what hierarchies or authority figures operate within the school - no authority role below that of headmaster being clear to them. Subsequently, perhaps because of the rather isolated work situation of the teacher behind the closed door of the classroom, the implications of departmental or pastoral hierarchies seem only very slowly to become apparent to the new teacher. Only sensitive questioning can gradually elicit the extent to which the head of department monitors or controls curriculum development, or the head of year advises about or checks up on the way in which tutorial periods are run. Often, of course, such supervisory support is not provided, and this is perhaps one reason why the young teacher, no longer in touch with his college supervisor, is only too glad to have the ear of the researcher. (Where collaborative analysis is used, the therapeutic element in the individual discussions which comprise the first phase of this research procedure is often prominent in work with young teachers). But the chief reason why such teachers are usually ready to

participate in collaborative research is probably that they are not conscious of an hierarchical structure weighing down upon them. This lack of recognition of organisational levels is sometimes found even in more experienced teachers. They seem reluctant to credit the complex relational networks in which, in a large school, all are involved, with any functions of authority or control. They may assure researchers 'we all work as a team', 'We decide things between us', or 'it's up to me what I do'. Traditionally, the teacher in the British educational system has a wide degree of personal discretion, and this apparent autonomy is much prized. But to the researcher working in a school it is apparent that teachers are certainly not autonomous in all aspects of their working life. As in any other organisation, considerations of seniority and the hierarchies of sub-systems within the main organisation constrain and delimit the contribution made by individual members, and this applies to their research participation as well. It is often the case that the researcher needs to be sensitive on the teacher's behalf to the extent to which the latter's participation will involve or interest other staff.

Liaison

Once agreement to participate in a research programme has been obtained from teaching staff, the question of administrative liaison between researchers and participants becomes of first importance. Most large schools have a 'link' man or woman who is generally responsible for school external relations, and it is not unlikely that the head teacher will nominate this person as liaison contact for the researchers. Nevertheless such a person will not necessarily prove an appropriate or effective intermediary between the research team and other staff who are going to participate in the field work. However senior the link person who is generally responsible for the school's external relations may be, if the research team are actually moving into work within with school it may soon be apparent that a more appropriate liaison person, and indeed a vital first contact, is the senior member of staff responsible for the day-to-day administration of the school. Such an administrative organiser is not likely to be satisfied with information that certain members of staff will be occupied with research interviews during free periods, and cannot be called upon to deputise for other teachers - he or she will emphasise the need for research arrangements to be controlled at the administrative level and will prove to be the vital link-person if research is to proceed at all. Being in close touch with a senior administrative figure provides the further advantage that researchers are less likely to be blocked off by more junior administrative personnel such as secretaries or administrative officers. But though a protective secretary is likely to take the line that no member of teaching staff has time to take on anything extra in the way of research interviews or meetings, she may prove an invaluable ally if research reaches the stage of needing access to records or other documentation, so the research team will be well advised to take any opportunity to arouse her interest and involve her in the project.

If research arrangements are negotiated through role holders who are themselves peripheral to the main functions of the school, the difficulties which arise are compounded by the inherent marginality of the research role itself, which has been well described by Rodman and Kolodny (11). If one marginal man must use another marginal man (12) as his link with 'line' personnel, not much is likely to be achieved, and the research team will do well to exercise tact in linking up with a more potent figure.

Effective Working Groups

In the case of the research procedure of collaborative analysis the question of what constitutes an effective working group of staff will become crucial at the stage when composite documents are being discussed. But any research procedure based on the clinical model and requiring the ongoing participation of practitioners is likely at some point to reach a similar stage of working together. For such discussions, researchers will find it desirable to bring together groups of staff who have more than an apparently similar ranking in common. In a large school, middle management ranking exists in a bewildering array - Heads of School or Building, Heads of House, Year Group Leaders, Senior House Tutors, Faculty Heads, Heads of Departments and so on. But the researcher needs to ensure that such staff do, in fact, define themselves as an appropriate working group and preferably that they already hold meetings together for other purposes than the research. Unless they do, it may well be the case that they perceive their status as very strongly differentiated even though this is not apparent to the outsider. This is particularly likely to be the case for heads of lower, middle and upper schools, where higher status is likely to be associated with those dealing with older pupils. Subjects too have their hierarchies - a fact that has not always been clearly recognised by whoever originally grouped them together in faculties (where these exist) - and the status of individual heads of department may be closely linked with that of their subject. A lack of status congruence does not preclude the bringing together of groups of staff, but it will certainly be a factor in the contributions which individuals are able to make in the group situation (13). Not all status gradations will be apparent to the researcher while individual work is being done. The only woman in a grade of apparent co-equals may contribute vivaciously in individual interviews, but fall completely silent in the group situation. Such occurences add a new dimension to the researcher's perception of organisational arrangements within the school.

The most favourable conditions for the elucidation of differing perceptions in joint discussions will exist when the researcher is able to bring together the group of staff who already work together in a close and fairly balanced colleague-relationship, within one of the sub-systems of the school. Although they already regularly meet together for other purposes than the research, their 'business' meetings are almost certainly experienced as frustratingly brief, because of the exigencies of the timetable, and are liable to interruption because of minor emergencies. Day-to-day colleague collaboration under such circumstances provides considerable motivation for participation in the extended processes of collaborative analysis or other research procedures in the clinical mode. There are likely to be many issues which group members have been wanting to have out with one another for a long while, yet there has never been time available. When it becomes possible to arrange extended group discussion (14), lively exchanges between colleagues are likely to ensue, in the liberating situation of a research process which takes for granted the legitimacy of differing perceptions. The cohesion deriving from well-established relationships between colleagues who work together at a joint task on a day-to-day basis inhibits the group from tearing itself apart. Despite the undoubted therapeutic element in the researchers' interaction with participants, it is not the writer's experience that dependency on the researcher is created nor that personal relationships between colleagues are impaired. Teacher-colleagues appear to be remarkably resilient in their debates with each other, and the marginal role of the researcher in the organisation is experienced as an advantage when phases of the research programme reach completion and teachers relax into a more informal brand of the 'shop talk' in which they have been avidly engaging as part of the collaborative research process. During such informal 'winding down' the researcher can withdraw without flurry. Kogan (15) has drawn attention to the need to guard against research programmes which focus on on-going organisational development acting as 'system-busters'. He suggests the need for a 'different kind of researcher' to carry out such work. Certainly the researcher enlisting teacher participation needs to bear in mind that working relationships within the school have to go on, even when the research is completed. His own sensitivity to the complexities of such relationships between teachers, most of whom have multiple work roles, will be an important factor in the completion of the research programme, and the willingness of teachers to repeat the participant research experience.

NOTES

1. B. Cave and C. Shroeder. 'The Teacher and Researcher'. NFER 1970. p.32. See also Chap. 6 (p.58) which points out that many teachers would like research to become less remote. Participative research as described in this paper meets many of the comments made to Cave and Schroeder by teachers about the feasibility of closer contact between researchers and teachers.

2. A.W. Gouldner, 'Explorations in Applied Social Science' in 'Applied Sociology - Opportunities and Problems'. Ed. A.W. Gouldner and S.M. Miller. Collier Macmillan, 1965.

3. A notable exception to this statement is provided by E Richardson's study, 'The Teacher, the School and the Task of Management', Heinemann, 1973.

4. For example, 'Schools, Parents and Social Services Project' carried out by the Educational Studies Unit, Brunel Institute of Organisation and Social Studies. The project focusses on pastoral care in secondary schools, the relationship between school and parents, and also between the school and other agencies having an interest in the welfare of young people in the secondary school age group. In the preparation of this paper, the writer has drawn on experience gained in the convening of this project.

5. B. Cave and C. Shroeder, op. cit. p.64.

6. For example, the Ford Teaching Project (carried out by the Centre for Applied Research in Education, University of East Anglia) is action research which enlists participation of teachers.

7. F.W. Neff, 'Survey Research: A Tool for Problem Diagnosis and Improvement in Organisation' in A.W. Gouldner and S.M. Miller, op. cit.

8. M. Kogan and W. van der Eyken, 'County Hall', Penguin 1973.

9. The masculine pronoun is used throughout this paper for the sake of brevity.

10. Few of the teachers participating in the 'Schools, Parents and Social Services Project' (see Note 4) had much experience of contacts between the school and the social services, or other social agencies. But almost all had some knowledge of contacts between the school and parents, and many were motivated by dissatisfaction regarding the effectiveness of all these contacts to take an initial interest in the project.

11. H. Rodman and R.L. Kolodny, 'The Researcher - Practitioner Relationship' in A.W. Gouldner and S.M. Miller, op. cit.

12. The role of the Counsellor in a large school is marginal almost by definition, yet he may well be responsible for many aspects of the school's external relations.

13. W.A. Westley in 'Dialogue with Reality' (forthcoming) discusses the implications of differing status both within research groups and in the researcher-participant relationship.

14. Because of the number of people involved, and the depth of the analysis undertaken, such meetings inevitably take place after the close of the school day. Teachers prove very willing to stop on after school, though the choice of day is often a difficult one.

15. Maurice Kogan, in an address to the Inaugural Meeting of the British Educational Research Association, 1974.

TEACHER PARTICIPATION IN A LONGITUDINAL STUDY

John Sayer.
Banbury School.

This brief note outlines in cryptic form two research projects initiated within Banbury School and sponsored by the D.E.S. Further information concerning the project may be obtained from the author.

Research Project I

To examine the differences, social and academic, which arise in the early years of secondary education from two systems of ability grouping, homogeneous and heterogeneous, within a single controlled situation.

Setting

Banbury School, taking all secondary pupils from Banbury and placing them in four largely separate "Halls" of matched intakes from the same primary schools. Each Hall takes approx. 4 forms of entry per year, age 11 for four years, taught on common curriculum but organised into different types of ability grouping. Pupils then move into the Upper School.

Organisation of Research Project

Funded 1971-1975 by D.E.S. through Oxford Department of Educational Studies in association with N.F.E.R. Steering Committee includes these three plus cross-section of Banbury School staff & L.E.A. and primary school members. Research Officer straddled between Oxford Dept., with computer programmer and secretarial services, and Banbury School, where part of management team and where one full-time assistant is based for classroom studies and individual interviews.

Coverage

Overall measurement of whole sample; emphasis on statistical analysis of quantifiable effects of groupings, inc. pupil questionnaires on attitude, opinion, and friendships, verbal and non-verbal tests, creativity tests, and subject-based achievement tests. Also comparison of teacher-pupil interaction and working methods in the two systems. Latterly, development towards qualitative examination of effects, inc. case studies.

Research Proposal II

To investigate what are the longer-term effects of different grouping systems in the early years of secondary education:

a) on pupil-teacher interaction and attitudes;
b) on academic and social development to age 16;
c) on levels of subject and career aspiration and on attitudes to subject and career "status" at the stages of 14-plus and 16-plus.

Organisation

Formal commitment from the school itself, needing support from partners in existing project, either as bridge towards involvement in research including other schools, or as full project in itself.

Further Information: see article in "The School as a Centre of Enquiry", Pubansco, Banbury School 1975.

The Banbury Research Project is of special interest
- because it is the only school-based project funded by D.E.S.
- because it arose from initiative of the school itself.
- because it examines an area of national interest & concern
- because it has involved teachers and become part of the school scene.

Teachers' expectations
- a detached enquiry to illuminate an area in which teachers' views in the school were irreconcilable
- reassurance
- grounds for decision
- professional enrichment

Teachers' involvement
- absorption of Research Officer and Assistant
- numerical dominance on Steering Committee
- administration of tests, individual interviews,
- openness to classroom observation

Problems of feedback to teachers
- distortion of research programme
- resisting demand for individual comparisons
- staffing changes and changes of professional climate.

Effects over four years
- early anxiety for findings thwarted
- undramatic results reassuring but unexciting
- watchdog element, acceptance replacing initiative
- enabled modifications to each community without polarised decisions
- led to questions beyond its own scope, and to demand for second project
- became a part of way of life.

Research Project II
- both to validate current project and to answer the queries which have become uppermost in the school.
- involves use of data from current project and using it in ways which must be of school's own initiative but also depend on specialised research support
- shifts emphasis to relationships and individuals as object of study, and shifts from synchronistic to diachronistic approach.
- sharpens the bluntest point of school management - assessment evaluation, consciousness of what is the effect of one's actions.
- appointment of assistant for research and statistics, and involvement of computer specialist from staff illustrating the shift towards school as centre of enquiry.

THE FUNCTION OF CLASSROOM OBSERVATION WITHIN THE METHODOLOGY OF THE FORD TEACHING PROJECT

Clem Adelman.
University of East Anglia.

The Ford Teaching Project involved the co-operation of the central team (John Elliott and Clem Adelman with Research Secretary Tina Reay) and forty teachers in twelve schools in East Anglia. The project was funded by the Ford Foundation who made no stipulations as to the direction of the research other than as set out in the proposal.

The project was sub-titled 'The Problems of Implementing Inquiry/Discovery Pedagogy within an Action-Research Framework'. This article has been requested to contribute an account of the methodology of the project. To me, methodology would make explicit the researcher's interpretative scheme - the process through which undifferentiated information is transformed into data. Research methods or techniques are used by researchers to assist or even, in unreflective cases, replace methodology. The Ford Teaching Project believed that:

1. In the process of becoming competent, a teacher develops a 'theory' of teaching.

2. That most teachers are capable of developing their pedagogy by systematically monitoring both the intended and unintended outcomes of their actions.

The procedures, methods, and methodology of the project stemmed from these two beliefs. To test their validity an action-research project was designed which basically involved the collaboration of the central team and the project teachers within a mutually acceptable ethical relationship.

We located teachers who aspired to be using inquiry/discovery pedagogy all or part of their teaching time. In most cases County Education Officers were initially contacted and they then generally delegated to education officers the task of locating school heads who reputedly had teachers using inquiry/discovery pedagogy within their school.

In the Autumn of 1972 John Elliott (the grant had not at that time been fully approved) visited all the schools recommended by the LEAs. In about half the schools the head, on being told of what inquiry/discovery pedagogy might involve, said "that sort of thing did not go on in his school". Maybe this is an instance of the reputation of some education officers as being out of touch with the grass roots. Whether this is the case or not, this may be an indication that the researcher cannot naively accept the conventional labels that are placed on schools by education administrators.

The initial conference of the Ford Teaching Project brought together thirty-eight teachers, three advisers, and one College of Education lecturer, and the central team. The central team resisted any demands to define what inquiry/discovery pedagogy might involve, even though this was demanded by a considerable number of the teachers. The central team felt that if they put forward their own opinions and definitions, not only would their statements be used as authoritative guidelines by teachers conducting the research, but also would inhibit the teachers from developing their own independent research. Much of the initial conference involved clarifying procedures and activities for the next term. The idea put forward by the central team that the teachers should monitor their own teaching by research methods devised by themselves was often greeted with alarm by the teachers. They felt that they should teach and the researchers should research; that they were not competent to research, etc. In the event, although the central team and the researching teachers were still differentiated by other aspects of their roles, the 'who does the research' boundary was eventually broken down in most cases. The reality was that both the central team and the teachers were researching, the teachers on their own teaching, and the central team attempting to monitor the teachers' monitoring overall.

The members of the school team met regularly to discuss and clarify immediate issues which were raised for discussion at regional meetings, compared to experiences of several school teams. The conferences consolidated and orientated for the future.

The teachers had decided by the second conference that the overriding feature of inquiry/discovery pedagogy was to enable independent reasoning. However, the pedagogic means to achieve this aim differed between teachers, although through eliciting meanings of the teachers' vocabulary for describing their teaching actions, the possible approaches to implementation of practice by teachers were made explicit. The central team suggested that to implement the aspired approach the teacher had to recognise some impediments to realisation, to devise and try experimental strategies for their resolution, and monitor the effects of these strategies to provide and suggest further problems and hypotheses. This whole developmental spiral was actualised through self-monitoring research techniques. These techniques were initially modifications or emulations of the techniques used by the central team with the teachers. Basically, these were observing and recording lessons and interviewing pupils. At the commencement of the teachers' own research, many of the problems to be recognised were, in fact, suggested by the central team only after some documents had been distributed by the central team illustrating the connection between the recognition of problems and hypotheses - for instance the two triangulation studies 'The Tins' and 'Three Points of View in the Classroom'. Most of the teachers began to do their own independent research based on their own research methods. These methods are reported in 'Ways of Doing Research in one's own Classroom'.

To fulfil the ethic of full collaboration, the central team fed back all research information, within the bounds of confidentiality, to individual teachers in the form of transcriptions of lessons and interviews. Although many of the teachers subsequently said that these were in retrospect the most potent means towards self-reflection, most head teachers found them "dull and boring" according to a brief independent evaluation by a visiting New Zealand educationalist.

The central team found as the project developed that they not only had to monitor the teachers' research, but also had to try to communicate with heads about what the teachers were doing in some cases. In addition, the central team had to consider the validity of their ways of interpreting and structuring what was going on in all the schools. The twelve schools were split into equal numbers based on geographical position around Norwich and Cambridge and half were visited by Adelman and half by Elliott. Consequently, the constructions of what was going on within the project were often argued out within the central team on the basis of their own experiences and then written up and offered for reality testing to the teachers in the form of circulated documents. Many of the teachers said that the documents were too long and the language incomprehensible. However, this criticism diminished as the project developed and clear lines of inquiry became to be shared by the teachers and the central team.

One of the most effective documents at this stage was only three pages long, containing a list of hypotheses for teachers to test. These hypotheses were derived from interviews with teachers and pupils, from observation, and from school and regional meetings. This document (J, November 1973) was greeted by many of the teachers as the first explicit guidance to them as to what they might be researching. The central team did not want to define the lines of research, as has already been mentioned, but teachers, as we all know, only have a limited amount of time and need some direction. By the time document J was circulated most teachers were already capable of healthy scepticism, which encouraged them to test these hypotheses in their own classrooms. Emerging from this were the beginnings of the development of their own hypotheses and in some cases case studies of their own classrooms.

Those teachers who could accurately monitor their own performance, that is their accounts seemed to be similar to the observer's and the pupils', were considered by the central team as reliable informants. The central team's problem, however, was to establish how to elicit honest accounts. Those teachers whose accounts did not match the observer's and the pupils' were considered to be in need of further practice self-monitoring the gap between their aspirations and their practice.

Reliable informants could be interviewed about problems and hypotheses being raised by other teachers. They also contributed considerably to the evolving of the common vocabulary to describe teaching actions involved in inquiry/discovery pedagogy and, incidentally, in other pedagogies. Of course we hoped that all teachers would become reliable informants eventually. However, misleading informants - if that is not too much of a value judgment - within the project were of as much value as the reliable informants, as they made clear to the central team that there were still considerable problems of understanding and still much progress to be made.

The teachers who effectively were monitoring the gap between aspiration and practice were not unduly perturbed by tape-slide recordings of their lessons. However, those that had not developed their self-monitoring were sometimes alarmed and surprised. These recordings were said by teachers to often pinpoint problems in a very revealing way.

Unlike the relative non-selectivity of tape recordings, the taking of slides is much more selective. Certain rules were followed in all classrooms for the taking of slides:

(1) When the teacher moves to a new group of pupils.

(2) When the teacher introduces a new object into the group or takes an object out of the group.

(3) The number of pupils focussed on the activity increases or decreases.

(4) The location of the group changes.

Within the limits of the technology, these four rules were sustained. The central team didn't use systematic observation schedules, although some of the teachers devised some suitable to their self-monitoring. John Elliott and I considered that as the project involved teacher development and to a large extent depended upon teachers testing their own theories about their practice, the central team should not specify any behavioural features to be assigned as characteristic inquiry/discovery pedagogy.

Now the research stage of the project has been completed, the documents do specify particular actions which the project as a whole has found to be relevant to both intended and unintended outcomes of actions. These are now in the form of hypotheses for teachers to test in the complexity of their own classrooms.

OBJECTIVITY, IDEOLOGY, AND TEACHER PARTICIPATION IN EDUCATIONAL RESEARCH

John Elliott.
University of East Anglia.

TEACHER PARTICIPATION IN THE FORD TEACHING PROJECT

As Clem Adelman has already pointed out, the Ford Teaching Project (1972-74) was an attempt to help teachers who aspired towards implementing inquiry/discovery approaches in their classrooms to reflect about the problems of realising these aspirations in practice. The central team (John Elliott and Clem Adelman) hoped that as a result of reflection upon their own and others practice, teachers would be able to identify problems, and formulate hypotheses about their causes, which were sufficiently generalisable to provide worthwhile orientations for any teacher who shared similar educational values.

The one problem we faced was that of developing a set of categories which would help teachers to classify their classroom data. We felt that the trouble with much educational research lies in the fact that the way classroom data is conceptualised is conditioned by the theoretical interests of the researcher. The result all too often is that such categories fail to generate knowledge which is relevant to the practical interests of teachers. So we decided to involve our teachers in the process of developing a conceptual framework for the study of classroom situations. We thought it important to investigate how they conceptualised and thought about classroom situations.

We did this via a study of the language they used when talking to each other about classrooms. Through language we hoped to get at the meanings it was being used to convey.

TOWARDS A TYPOLOGY OF MEANINGS

Our first project conference involved detailed discussions of recordings and transcripts of classroom situations. It was noticeable at the time that teachers had severe communication problems. Some appeared to be using similar terms to convey different meanings while others used different terms to convey similar meanings. So after the conference we studied the taped discussions with a view to discovering terms that frequently recurred. The terms used with the greatest frequency were:

subject centred
child centred
structured
unstructured
framework
open ended
guided
directed
dependent (child)
independent (child)
formal
informal

We then probed the meaning of these terms in interviews and discussions with teachers. The results suggested that the following were interchangeable:

'structured' and 'framework'
'structured' and 'subject centred'
'unstructured' and 'child centred'

There was a high degree of interchangeability between 'formal' and 'independent' and 'informal' and 'independent'. However, 'formal/informal' was sometimes used to contrast social organisation in the classroom. Here 'formal' picked out a centralised classroom in which teachers could monitor the words and actions of all the pupils at a glance. 'Informal' picked out a decentralised classroom in which teachers were not able to monitor everything that was going on at the same time.

So the most frequently recurring terms were being used to make three main contrasts. In order to facilitate communication we suggested that only one set of terms should be used to pick out a particular contrast. We recommended the following:

formal - informal
structured - unstructured
directed - guided - open ended

'Formal - informal' picked out one dimension along which teacher influence in the classroom might vary. The more the teacher's actions prevented pupils from reasoning independently, the more formal the situation, and the more they left the pupil free in this respect the more informal the situation.

'Structured - unstructured' picked out a dimension along which the teacher's aims might vary. The more aims were primarily, but not exclusively, concerned with preselected learning content, the more structured the teaching. The more aims were primarily concerned with protecting and fostering independent reasoning - with the manner rather than the content of learning - the more unstructured the teaching. A structured approach does not necessarily exclude independent reasoning as an aim; its protection or fostering may be aimed at within an overall structured approach.

'Directed - guided - open ended' picked out a dimension along which the teacher's methods might vary. Directed methods indicate a high degree of attempted control over pupils' thinking. Guided methods indicate a lower degree of attempted control over pupils' thinking. Guided methods indicate a lower degree of attempted control. The teacher tries to exert some positive influence on pupils' thought processes but in doing so is concerned to protect or foster their powers of reasoning. The teacher 'guides' when he tries to respond helpfully to problems posed by pupils in the learning situation. Open ended methods indicate a low degree of attempted control. They constitute attempts to remove situational constraints on independent reasoning rather than attempts to exercise any

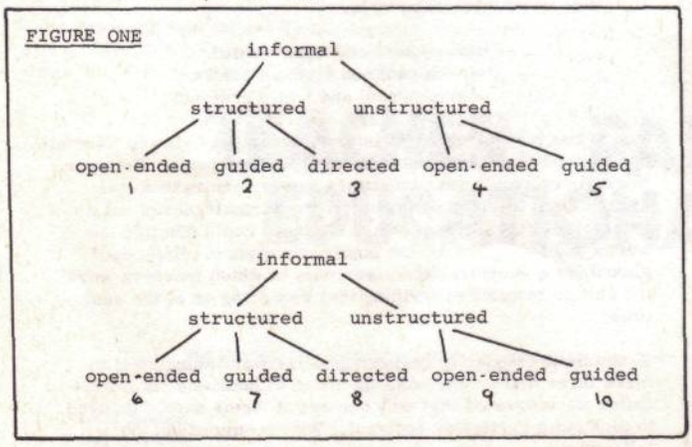
positive influence on intellectual development.

The three dimensions can be represented as follows:

1. Situation — formal - informal
2. Aims — structured - unstructured
3. Methods — directed - guided - open ended

It is obvious that these categories, elicited from teachers, are heavily conditioned by the importance they ascribe to independent reasoning in the classroom. If the situation is a formal one, independent reasoning is not possible. Its possibility rests on the creation of an informal situation. The distinction between structured and unstructured aims is also relevant in this respect. The extent to which the teacher has strong preconceived ideas about what is to be learned may influence the extent to which an informal situation can be created and independent reasoning fostered within it, as may the degree of control he tries to impose on pupils' thinking. The categories are a good illustration of the way teachers conceptualisations of their practice are conditioned by their practical concerns or values. They pick out significant variables whose interaction may influence the extent to which independent reasoning is protected and fostered. The possible patterns of interaction between these variables can be illustrated by combining categories from each dimension to produce a typology of teaching:

FIGURE ONE



The typology takes into account both intended and unintended consequences of teaching. Types 3, 6, 7, 9 and 10 involve unintended consequences, while 1, 2, 4, 5 and 8 involve intended consequences. Take 2 and 7 as examples of each group. In 2 'structured - guided' indicates an attempt to pursue preconceived learning content while trying also not to impose constraints on pupils' freedom to reason this content out for themselves. The fact that an informal situation was maintained indicates that the pursuit of structured aims did not impose constraints on independent reasoning. In 7 the 'formal' category indicates that the way structured aims are pursued create a formal situation in spite of the teacher's attempt to refrain from controlling pupils' thinking too rigidly.

Thus we can see how the categories help teachers to monitor their own and other people's practice in the light of their practical concern to protect and foster independent reasoning. The typology outlined was used throughout the project by central team and teachers as the conceptual basis for the analysis of teacher behaviour in classroom situations.

ON PRODUCING ACCOUNTS

Another problem we faced was that of helping teachers to produce accurate accounts of both their actions and the effects of their actions on pupil performance in the classroom. We implemented a triangulation procedure in which the teacher formulates a final account through dialogue with his puils and an outside observer. Each party in the triangle is in a special epistemological position. The teacher has direct access to the intentions embodied in his actions through introspection.

His pupils are in a good position to assess the causal significance of his actions for their own performance, because they have direct access to the subjective meaning of their responses and to the interpretations on which these responses are based.

Teachers' actions elicit responses because pupils interpret them as indications of the responses expected. In other words teachers try to elicit responses by indicating in their behaviour what they want. Their accounts of the causal significance of their actions always rest on beliefs about how pupils have interpreted them. If these beliefs are incorrect then their inferences about the subjective meaning of pupils' responses are unlikely to be correct. For example, a teacher tries to discover whether pupils agree with his views on a certain issue by saying "Do you all agree?"; the pupils say "yes", which is interpreted by the teacher as an expression of agreement. His interpretation of pupil response here assumes that they understood his action as 'an attempt to discover whether they agreed'. But suppose they in fact understood it as an 'invitation to agree with him' instead. In this case a more intelligible interpretation of their response would be 'they pretended to agree with him'. How pupils understand the teacher's actions places certain logical restrictions on what can count as an intelligible response to them. Therefore, in order to check his account of the causal significance of his actions teachers need to look not only at the evidence of observable pupil behaviour, but also at pupils' accounts of what their responses mean and the interpretations on which they are based.

The outsider is also in a special position with respect to observable evidence of teacher/pupil interaction by virtue of his detachment from the action.

In the light of these considerations we felt that it was important for teachers to enter into dialogue with pupils and outside observers as part of the process of trying to produce objective accounts of their practice. The triangulation procedure we initially adopted involved the observer eliciting a retrospective account of a lesson from the teacher. The observer then interviewed the teacher's pupils with his permission and negotiated with them the teacher's access to their accounts. He then made available to the teacher his own, the pupils, and the teacher's accounts to compare and contrast. After the teacher had had an opportunity to reflect on the material, the teacher and his pupils, plus the observer, when it was possible, discussed discrepancies in their respective accounts following a tape or tape-slide recording of the lesson or part of it. The teacher was then in a position to formulate a final account in the light of the discussion.

Initially in the project the observer, who was usually a central team researcher, took a leadership role within the triangulation. He elicited the teacher and pupil accounts and negotiated access to each other's accounts. Teachers found it extremely difficult to elicit honest accounts from pupils without the help of the observer at first. However, if pupils discovered that the teacher was able to listen to their views and consider them reflectively then teachers found it easier to elicit honest accounts without help from the observer. As the project progressed many teachers began to adopt a leadership role in the triangulation. Not only did they begin to initiate open dialogue with their pupils, but were able to produce detailed accounts of their practice without these having to be elicited by the observer. They also developed a more independent relationship to the observer, using him to help them to monitor specific aspects of their performance which caused them concern.

The triangulation procedure I have described enables three parties concerned with the truth about a situation to construct objective accounts by drawing on each other's special perspective, and exploring discrepancies which arise.

The reader might like to compare and contrast the following excerpts from a triangulation study of an environmental studies lesson:

EXTRACT FROM INTERVIEW WITH TEACHER.

Observer	Do you know that you use the words "Do we all agree?" quite a lot?
Teacher	No I didn't (pause) OK - I know I use that a lot.
Observer	Three or four times.
Teacher	I am asking for assent.
Observer	Are you? Is that what you are asking?
Teacher	I think probably I am. I think possibly I use that when I don't get - - if I make a statement and I haven't got a - - I don't know sometimes if it is

a rhetorical question or whether it is a question I want an answer to or whether it is just a statement, but I make a statement and I hope the response will come from it. If a response doesn't come from it, you either repeat it in a different way to a single individual and put them on the spot, or you perhaps get over it by saying OK or 'Do you all agree with that?'. I suppose they can possibly con me by saying yes and carry on. It is something I hadn't thought of.

Observer — Do they all say yes?

Teacher — Well they didn't all say no. I reckon if you take a non-negative approach to be an affirmative, which is perhaps a big thing to do. I don't think you ought to do that really. Yes that's naughty isn't it!

Observer — Well the thing is I suppose when you say "Do we all agree?" they can say no.

Teacher — I give them the opportunity to say no.

Observer — They can say no, but how do they see it. If they see it as your seeking agreement - -

Teacher — I think a lot of the time one must be seeking agreement . . . what I am trying to put forward is what I feel to be a reasonable statement; a true statement. Although I didn't today, I do in fact sometimes put forward daft statements and you do usually find that they disagree if there is something stupid. It was a bit tame today - I mean you were coming in part way through a situation which wanted finishing and therefore I finished it. In terms of them going away and doing things - and I thought you would be more interested in discussion because of the material you had got - your recording technique . . .

Extract from observer's notes (written during lesson).

Look at old tables of results. What's happened? Teacher asks specific pupils questions. When he disagrees raises his voice quizzically as if he disagrees. Question and answer. Hints. When right answer is given it is reinforced by the teacher. 'Right'. (guessing game) Do you all agree with that? Reply by one boy 'Mm'. When boy responds in a way which doesn't fit what teacher wants it is chopped. People not encouraged to elaborate on ideas. Wants to get them critical of John Innes compost manufacturer. John Innes made by pupils promotes growth better than commercial product. Asks why paper pots are better than plastic pots. Often makes an interpretation. Asks pupils if it is a 'reasonable guess'. Someone murmurs again 'yes'.

Interview with Pupils

Pupil — But he wouldn't ask you what you think your conclusions were, he'll put his own conclusion up on the board, and you have to write it. He says do you agree, not always but he don't want to rub if off so you just say yes to keep him quiet.

Observer — You say yes to keep him quiet?

Pupil — Keep him happy.

. . .

Observer — There was a time when he said he was making a guess and he asked you if you agreed whether it was a reasonable guess. I don't know if you remember that?

Pupils — Yes.

Observer — And one person said yes and everybody else kept quiet. Now what I want to know is whether the person who said yes really did agree with him or just said yes because they thought he wanted them to say yes, and why everybody else kept quiet?

Pupil — Well he would have liked us to say yes, really, cause I mean you could see it.

Pupil — If you'd said no you'd waste time arguing wouldn't you.

Pupil — Yeh, if you ever say no he'll stand there and just keep on and on.

Pupil — He'll keep on till you come to his way of thinking.

Pupil — So it's best to say yes to start with.

Observer — So even if you did disagree when he said "Do you all agree?" you wouldn't.

Pupil — If you said no he'd keep on to you until you said yes.

Pupil — If you said no he's going to say why not.

Pupil — And if you argued with him he'd come round to the same point where you left off.

Pupil — Back to his way of thinking.

Excerpt from tape-recorded lesson.

Teacher — Yeh, do you all agree with that?

Pupil — Mm.

Teacher — What do you think Derek? I mean are you bothered?

Teacher — . . . Would that be the only thing you want to know about a plant?

Pupil — How to condition it Sir.

Teacher — Yeh, do you all agree with that?

MICROTEACHING

SELF-CONFRONTATION REVISITED

Gordon MacLeod.
University of Stirling.

PREFACE

'One cannot read the thoughtful review of research on self-confrontation techniques by Fuller and Manning without coming to the conclusinn that these self-confrontation techniques are filled with recognized hazards and unknown pitfalls. Such technological approaches to role training should be studied, but always by professionals capable of handling the serious mental and behavioral disturbances that the techniques are known to be capable of producing. Individuals who see thems elves in live action on a screen respond emotionally to what they see. Typically they become preoccupied with the appearance of their body and fail to take note of the behavior they see and the consequences of that behavior.' (Travers, 1975).

INTRODUCTION

Microteaching is now a recognized aspect of teacher education, and its very popularity means that it is probably the largest single source of so-called 'self-confrontation' experiences in teacher education. That this is so, is not a consequence of the original purposes and principles of microteaching, for as McDonald (1973) notes, 'Originally, microteaching was devised as a procedure for facilitating behavioural control' (p. 71), although, as its popularity grew, 'many users of microteaching did not see the relevance of behaviour modification principles' (p. 72) and 'the most undesirable consequence of the promotion of microteaching was that the role of behaviour modification in training was obscured' (p. 73).

As McDonald has detected a change in the principles underlying microteaching, so is it possible to discern what seem to be changes in microteaching practice, and in particular a move towards the almost universal use of videotape feedback. Thus, for example, writers on the early Stanford microteaching programmes suggest that although the wedding of videotape and microteaching was a happy association, it is by no means a necessary one (Allen and Ryan, 1969). By contrast, later writers such as Foster, Heys and Harvey assert that the use of video feedback 'is now the dominant feature about which (microteaching) programmes are invariably organized' (p.100), a view which is echoed by Griffiths (1972).

A possible reason for the popularity of videotape feedback is that its effectiveness is well attested in reviews of the research literature, despite the fact that the literature does not undeniably support this conclusion. Stones and Morris (1972) are therefore probably correct in suggesting that 'it is generally agreed . . . that the availability of video recording enhances the flexibility and effectiveness of microteaching' (p. 82), in that such agreement can be found among Borg and his colleagues (1970), Perlberg (1970), McAleese (1973) and Turney and his colleagues (1973). More recently, a few polite murmurs of dissent as to the efficacy of videotape feedback have been apparent (Griffiths, 1974, Brown, 1975). However, for the purposes of the argument being advanced here, it is not the validity of these statements which is of concern; rather is it their near unanimity in support of the effectiveness of video feedback.

A further mark of the increased importance of videotape feedback was the publication in 1973 and 1974 of two important review papers on self-confrontation (Fuller and Manning, 1973; Bierschenk, 1974). The former is particularly well known, and its publication was well received in this country, with the editor of this journal noting that the 'original paper... so impressed the editorial board that they wished to bring it to the attention of readers as soon as possible' (British Journal of Teacher Education, 1975, p. 97) and this was achieved through a short report on the paper by Finlayson (1975).

What follows in this article will be generally critical of Fuller and Manning, but this is not intended to detract from the value of their review. Indeed, many of the difficulties encountered with their review are recognized by the authors - for example, their explicit statement that their concern is 'whether teacher education video playback practices can be informed by the experimental and theoretical literature on confrontation, particularly that literature outside teacher education' (p.472). What is perhaps grievous is that by making use of this literature from outside teacher education, particularly that of psychotherapy and counselling, the authors have stretched some conceptual definitions considerably, to include very different operational definitions, although, once again, the authors do recognize this and admit that 'we took some liberties with construct names in order to classify them.' (p. 473)

STRESS OUTCOMES

It is the grouping together of studies using different procedures and different measuring instruments which is the nightmare of all reviewers, and it is in this area that some violence has been done to what ought to be our conceptualization of video playback in teacher education. For example, one of the studies grouped by Fuller and Manning under 'Stress Outcomes' of self-confrontation is that by Nielsen (1962), an oft-quoted study to show that self-viewing produces stress. However, few reviewers note that this study was deliberately designed to produce stress in the subjects before self-viewing took place. A reading of the extended title of this book reveals a study of 'Viewing a sound motion picture of self and another in a stressful dyadic interaction'. Also quoted under the heading of 'Stress Outcomes' is Salomon and McDonald's (1970) finding that 'Most self-observations were negatively evaluated by the viewers', a finding which would seem to be just as consistent with low stress as with high stress. Indeed, evidence from research at Stirling not only replicates this finding of majority of negative evaluations, but also suggests that this negativity is perhaps a reflection of realism about the viewed performance. That is, when a measure of negativity obtained from 49 self-viewers in microteaching is correlated with their performance and their estimate of performance on the viewed lesson, the resultant pattern of significant relationships suggests that those whose comments are most negative are those who rate their overall success as low, who estimate their performance on specific behaviours to be 'poor', and whose actual performance is observed to be 'poor' (Table 1). The skill being practised by these students was the FIAC-based 'Pupil Initiation -Teacher Response', and the operational definitions of these concepts were in terms of nine of the ten basic Flanders' categories.

TABLE 1

Selected Correlates of Overall Negative Evaluation

(Second Lesson)

Proportion of Statements Coded as Negative with	r	p
Rating of Overall Success	-.64	<.001
Proportion of 9's on Teach	-.56	<.001
Estimate of 9's on Teach	-.56	<.001
Proportion of 8's on Teach	.33	< .05
Estimate of 5 + 6 + 7 on Teach	.29	< .05

Although the findings shown in Table 1 are by no means conclusive, they do allow speculation that the negativity of evaluation is based on a fairly realistic self-assessment of performance rather than being a reflection of stress.

That stress outcomes of self-viewing may have been somewhat exaggerated is further suggested by a study at Stirling by Gilmore (1975) involving in-depth interviewing of thirty-eight student teachers who had completed a semester-programme of microteaching. Of these thirty-eight, only one found self-viewing 'definitely upsetting', while another thirty found it definitely 'not upsetting'.

Further circumstantial evidence, using a very different measuring instrument, is that provided by McAleese (1975). By means of heart-rate measurement on university lecturers, he found that self-viewing did produce a marked increase in heart-rate to about 110 per minute at the beginning of self-viewing, declining to about 100 per minute at the end of the twelve-minute self-viewing period. Whilst these raised levels may indicate a moderate stress situation, they by no means compare with the raised level of heart-rate while lecturing, particularly that occurring when questions are asked of the lecturer where peaks of 140 and 150 per minute were observed.

Thus, while it may be the case that self-viewing produces moderate stress (perhaps better named arousal), it is by no means apparent that this stress is any more than that produced by any other 'public' appearances.

FOCUS ON BODY AND VOICE

Further doubts about the emphases conveyed by Fuller and Manning concern their section 'Focus on Body and Voice'. They write that 'Unless some other focus, perhaps some powerful focus, is provided, self-viewers seem to focus on themselves' (p. 474). The first study they cite is that by Lawrence (1971) who found that the self-viewers he studied were difficult to dissuade from their focus on negative body characteristics. However, the self-viewers he studied were nude marathon group members.

The two studies which Fuller and Manning do cite which seem most directly to concern teacher education are these by Salomon and McDonald (1970) and by Bedics and Webb (1971). In the first of these studies, no guidance was given to subjects as to how they were supposed to react to the self-viewing experience, but given that the subjects had just completed a structured questionnaire in which they had rated the eight concepts shown in Table 2 on nine seven-point rating scales, it is perhaps not surprising that the subsequent median percentage of 'physical appearance' observations was 58% whilst that for reported 'teaching behaviour' was 18%.

TABLE 2

Concepts used by Salomon and McDonald (1970)

Teacher education	Me as friend
Use of electronic devices in classroom	Me as teacher
Camera in classroom	My appearance in classroom
Me as student	Me as seen by students

The second study, by Bedics and Webb (1971) is reported as demonstrating that 'Both single and triple playback groups without training focused on themselves' (p.475), although it is also acknowledged that the triple playback group did focus more on the teaching act. Table 3 shows the Bedics and Webb date re-cast in a form relevant to this issue for the untrained group.

It is apparent from Table 3 that although it may be strictly correct to assert that there is focus on self after the third viewing, the data when in this form do change the emphasis considerably and suggest that focus on body and voice may not be quite so lasting as Fuller and Manning suggest, and indeed that even at first viewing, the maximum number mentioning any of the 'self' criteria represents only 50% of the group.

Content-analysed free-response data collected at Stirling during a microteaching programme, and shown in Figure 1, reveals the extent of 'Personal' comments after first, second and fourth microteaching lessons, as compared with the extent of comments on pupil behaviour and on teacher behaviour.

TABLE 3

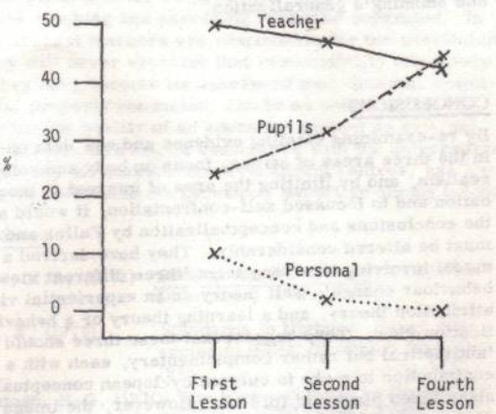
Number of Subjects Mentioning 'Personal' Criteria after First and Third Self-Viewing (derived from Bedics and Webb, *op. cit.*)

Criterion	First self-viewing (N = 22)	Third self-viewing (N = 22)
Voice	10	4
Appearance	4	0
Non-verbal mannerisms	6	1
Verbal mannerisms	11	3

Figure 1 indicates that in this focus-oriented form of self-viewing (which is much the most common in teacher education), those comments about observed behaviour coded as Personal, including body and voice, account for 10% of comments at first viewing, 2% at second viewing and 0% at third viewing

FIGURE 1

Percentage of 80 Foci in Selected Categories



(i.e. fourth lesson), whilst comments on teacher behaviour decrease from 50% of the total to 43%, and comments on pupils increase from 24% through 32% to 45%. (These are percentages of all observed foci on which comments were made and account for approximately 95% of all foci).

REALISM ABOUT SELF

Moving on to 'Realism about Self', Fuller and Manning note that 'It comes as no surprise to find that estimates by student teachers. . . about their own behaviour are very inaccurate (Johnston, 1969)' (p. 476). Unfortunately, Johnston, in the paper referred to, does not report his data but says 'Estimates by student teachers of the percentage of indirect teaching they exhibit in their lessons are very inaccurate under both traditional supervision and self-supervision' (p. 6). Data collected at Stirling does serve to illuminate this issue. Table 4 shows students' estimates of performance on five criterion measures hypothesised to be under their control, together with the measure of actual performance, the difference between actual and estimate, and the product-moment correlations between actual and estimate.

The results of Table 4 would seem to indicate a general bias in estimating one's own performance to be somewhat worse than it is, whilst pupil performance (8's and 9's) seems to be estimated with a high degree of accuracy. The significant and highly significant correlations between estimates and actual performance would seem to support the hypothesis that a fairly systematic error is being made in terms of 'being on the

TABLE 4

Comparison of Estimated and Actual Proportions of Criterion Behaviours (Second Lesson)

Criterion	(1) Estimate	(2) Actual	Difference	r_{12}	p
Proportion of:					
1 + 2 + 3 (Teacher Response)	12.94	24.83	-11.89	.35	< .01
4 (Teacher Questions)	21.21	16.13	+ 5.08	.47	<.001
5 + 6 + 7 (Teacher Initiation)	22.42	17.23	+ 5.19	.72	<.001
8 (Pupil Response)	18.70	16.38	+ 2.32	.47	<.001
9 (Pupil Initiation)	20.91	20.59	+ 0.32	.77	<.001

safe side' and devaluing one's own performance. It is worthy of note that the measure of indirect teaching used here (1 + 2 + 3) is very much the least accurate, and that this does seem consistent with Johnston's conclusion but not with Fuller and Manning's generalization.

CONCLUSIONS

By re-examining existing evidence and new data on self-viewing in the three areas of stress, focus on body and voice, and realism, and by limiting the area of interest to teacher education and to focussed self-confrontation, it would appear that the conclusions and conceptualization by Fuller and Manning must be altered considerably. They have derived a conceptual model involving what they term 'three different views of behaviour change': self theory or an experiential view, attribution theory, and a learning theory or a behaviour modification view. They suggest that these three should not be 'antithetical but rather complementary, each with a different contribution to make to current cyclopean conceptualizations about video playback' (p. 505). However, the image of self-viewer which they create in their review particularly in the sections on stress, self-focus and realism, may be parodied as that of student-teacher as an anxious, narcissistic and unrealistic victim of circumstances, being aroused and impelled to action by dissonance and blows to self-esteem, such motivation leading the self-viewer (in focussed self-confrontation) to model selectively-reinforced behaviours. The intention of this paper is to move this image towards that of self-viewer in teacher education as a rational and internally consistent processor of information. After all, trainees in teacher education are not psychiatric patients, are not counselees, nor, often, are they nude marathon group members. Fuller and Manning's assignation of 'causal status to cognitive variables' (Finlayson, op. cit. p. 102) is to be welcomed, but so is Finlayson's reminder that they have perhaps not gone far enough in considering the causal role of cognitive variables. Indeed it seems fair to assert that far too little attention has been given to cognitive variables, both as determinants of self-viewing performance, and as causes of subsequent behaviour. Given such a view, the theoretical perspectives we adopt must borrow much more from empirical and speculative work in social and cognitive psychology. In particular, variants of Kelly's (1955) personal construct theory, the notions of schemata, advanced reluctantly in these terms by Bartlett (1932), enthusiastically by the Piagetians (Flavell, 1963) and the notions of 'casual schemata' used by the attribution theorists (Jones et al., 1971), would all seem profitable in beginning to conceptualize what occurs during teacher self-viewing.

REFERENCES

Allen, D.W. and Ryan, K.A., (1969) Microteaching. Reading, Mass: Addison-Wesley.

Bartlett, F.C., (1932) Remembering. Cambridge: Cambridge University Press.

Bedics, R.A. and Webb, Jeaninne N., (1971) Measuring the Self-Evaluation of Teaching Behaviour through the Use of Videotape. EDO51 079.

Bierschenk, B., (1974) Perceptual, evaluative and behavioral changes through externally mediated self-confrontation. Didakometry, 41. Malmo, Sweden: School of Education.

Borg, W.R., Kelley, M.L., Langer, P. and Gall, M., (1970) The Minicourse: A Microteaching Approach to Teacher Education. London: Collier-Macmillan.

British Journal of Teacher Education (1975) Volume, 1 No. 1.

Brown, G.A., (1975) Microteaching: Research and Developments, in G. Chanan and S. Delamont (Eds) Frontiers of Classroom Research. Slough: N.F.E.R.

Finlayson, D., (1975) Self-Confrontation: A Broader Conceptual Base? British Journal of Teacher Education, 1, 1, pp. 97-103.

Flavell, J.H., (1963) The developmental psychology of Jean Piaget. Princeton, New Jersey: Van Nostrand, 1963.

Foster, J.K., Heys, T.A. and Harvey, J.M., (1973) Microteaching: A Review and a Study of the Effect of Microteaching on Teaching Effectiveness as Measured by Pupil Achievement. The Forum of Education, XXXII, 2, pp. 100-141.

Fuller, Frances F and Manning, B.A., (1973) Self-Confrontation Reviewed: A Conceptualization for Video Playback in Teacher Education. Review of Educational Research, 43, 4, pp. 469-528.

Gilmore, S., (1975) Student Reactions to Self-Viewing on Videotape. Personal Communication.

Griffiths, R., (1972) Some Troublesome Aspects of Microteaching. Paper presented to the Annual Conference of the National Educational Closed Circuit Television Association. University of Stirling, (Mimeo).

Griffiths, R., (1974) The Contribution of Feedback to Microteaching Technique in A.J. Trott (Ed.) Microteaching Conference Papers, APLET Occasional Publication, No. 3., pp. 15-22.

Johnston, D.P., (1969) The relationship of self-supervision to change in selected attitudes and behaviors of secondary student teachers. Paper presented at the American Educational Research Association, Los Angeles, California. (ED 029 823).

Jones, E.E., Kanouse, D.E., Kelley, H.H., Nisbett, R.E., Valins, S. and Weiner, B., (1971) Attribution: Perceiving the Causes of Behavior. Morristown, New Jersey: General Learning Press.

Kelly, G.A., (1955) The Psychology of Personal Constructs, Norton.

Lawrence, S.B., (1971) From zazen to videotape: Encounter group innovations. In L. Blank, G.B. Gottsegen and M.G. Gottsegen (Eds). Confrontation: Encounter in self and interpersonal awareness. New York: The Macmillan Co.

McAleese, W.R., (1973) Microteaching: a new tool in the training of teachers. Educational Review, 25, 2, pp. 131-142.

McAleese, W.R., (1975) Personal communication.

McDonald, F.J., (1973) Behavior Modification in Teacher Education, in C.E. Thoresen (Ed), Behavior Modification in Education. N.S.S.E. Yearbook, Chicago, Illinois: University of Chicago Press.

Nielsen, G., (1962) Studies in Self-Confrontation: Viewing a sound motion picture of self and another person in a stressful dyadic interaction. Munksgaard, Copenhagen: Scandinavian University Books.

Perlberg, A., (1970) Microteaching: a new procedure to improve teaching and training. Journal of Educational Psychology, 1, 1, pp. 35-43.

Salomon, G. and McDonald, F.J., (1969) Pre-and post-test reactions to self-viewing one's teaching performance on videotape. R. and D. Memorandum, No. 44, Stanford Center for Research and Development in Teaching.

Stones, E. and Morris, S., (1972) Teaching Practice: Problems and Perspectives. London: Methuen.

Travers, R.M.W., (1975) Empirically Based Teacher Education. The Educational Forum, 39, 4, pp. 417-433.

Turney, C., Clift, J.C., Dunkin, M.J. and Traill, R.D., (1973) Microteaching: Research, Theory and Practice, Sydney: Sydney University Press.

THE PROCESSES OF SUPERVISION IN MICROTEACHING

George Brown, University of Nottingham.
Brian McGarvey, University of Ulster.

THE PROCESSES OF SUPERVISION IN MICROTEACHING

This paper describes the development of a category instrument which may be used to analyse the processes of supervision in microteaching. It is not proposed to review the literature on supervision in teaching and microteaching (See Parry and Gibbs, 1974) or to comment upon so called theories of supervision (See Dussault, 1970). The opportunity to do so is tempting but that delicious task is left for another occasion. Instead it is proposed to make a brief comment upon the supervisory process and plunge into a description of the development of the instrument.

Supervision is a teaching situation and as such supervisors and students bring to it their previous experiences, attitudes, role perceptions and expectations. These are manifest in their interactions and hence we should look first at their interactions rather than search for personality characteristics which may or may not be salient in the supervisory situation. Every teaching situation is intended to yield changes in the learner. Hence in supervision we should look also at the outcomes of the supervisory process: changes in the teaching styles or skills of the student (Darr, 1973). A necessary preliminary to this endeavour is to develop a reliable and sensitive instrument for observing the processes of supervisor-student interaction. As you will see the task is not easy nor at this stage can we be sure that we have captured the essential features of the supervisory process.

THE MAKING OF THE INSTRUMENT

Stage 1: The search for categories.

Supervisory conferences at Ulster consist of four students and one supervisor. The students teach their microclass of pupils consecutively, observe each other's videotaped lessons as well as their own and discuss them and the next lessons with the supervisor. About 200 students are involved in microteaching each week. At the outset it was therefore decided to develop an instrument which would take account of the group's contributions and which would enable computer based analyses to be carried out on the data.

The literature search yielded four main but overlapping dimensions of the supervisory process. These may be labelled the Verbal Interaction Dimension, the Problem Solving Dimension, the Content Dimension and the Relationship Building Dimension. The Verbal Interaction Dimension describes the semantic structure of the discourse, the Problem Solving Dimension the verbal processes of the group, the Content Dimension the subject matter under discussion and the Relationship Building Dimension may be regarded as an attempt to classify the interpersonal relationships within a group. Fig. 1 sets out a summary of the possible components of the system. Flanders' (1960) and Brown's (1975) systems gave the foundations of the interaction system; Brown and Hoffman's (1966) and Blumberg's (1970) systems helped in categorising problem solving and relationship building behaviour and Heidelbach's (1968) system gave ideas on classifying subject content. Weller's (1971) full account of his system and code rules helped in the development process. The components were not defined rigorously at this stage. There were considerable overlaps in meanings both within and across dimensions and these were reduced or eliminated in the next stage.

FIGURE I: VARIABLES CONSIDERED

Verbal Interaction Dimension

tutor tells
tutor questions
tutor accepts, responds
student no. 1 responds
student no. 1 initiates
student no. 2 responds
student no.2 initiates
student no. 3 responds
student no. 3 initiates
student no. 4 responds
student no. 4 initiates

Problem Solving Dimension

managing
controlling
leading
structuring
informing
describing
identifying
defining
planning
completing
comparing
contrasting
explaining
opining
analysing
evaluating
reinforcing
praising
criticising
suggesting
hypothesising
advising
soliciting
questioning
reacting
responding
summarising
silence - comfort
silence - distress
silence - hesitation
levels of thinking

Content Dimension

structuring
teaching skill under focus
other teaching skills
lesson organisation, structure
lesson subject content
lesson materials, aids
pupils' behaviour
discipline and control
next week's lesson
next week's skill
microteaching
social emotional talk
other talk

Relationship Building Dimension

rapport building behaviours

praise
support
relief
of tension
laughter
empathy
showing
interest
acceptance
understanding
honesty in
appraisal
encouragement
frankness
conveying a
feeling of
co-operation
silence - comfort

rapport disrupting behaviours

criticism
rejection
hostility
cynicism
silence - distress

other behaviours

student autonomy
student dependence upon supervision

semantic rating scale (McCroskey and Wright, 1971)

tense - relaxed
flexible - inflexible
relevent - irrelevent
interested - apathetic
brief - lengthy

Stage 2: The clarification and collapse of categories.

Stage 2 consisted of observing a wide variety of supervisory conferences and using this experience to select and refine the categories of the instrument and to develop rules of coding.

Forty conferences were recorded involving 15 supervisors and 56 students drawn from different curriculum areas. Thirty recordings were made in Weeks 6 and 9 and ten in Week 11 of the microteaching programme. A full analysis of these recordings is being carried out during this academic year 1975-76.

At the outset it was clear that at least two sets of observations would have to be made on each supervisory conference. It also became clear that the Relationship Building Dimension

did not yield to relatively low inference-high observability measuring instruments. It was therefore reluctantly dropped from the system. The main criteria for selecting a category were that it was readily identifiable within the supervisory conferences, it occurred relatively frequently and it could be defined operationally. The observation instrument evolved through five phases of development. Each phase consisted of defining, ordering, observing and taking reliability measures. It is not proposed to describe these phases in detail. But it is worth noting that coding all the contributions of every student proved a vexing and unreliable exercise. Hence only the verbal behaviour of the supervisor is categorised in detail. This reflects existing supervisory practices - supervisors do talk most of the time. The present form of the category system is summarised in Figs. 2a and 2b. 2a gives the categories to be used in the first session of observing and 2b those to be used in the second session. Rules for coding were developed and the present set may be obtained from Dr. Brian McGarvey.

Figure 2 A
THE MICROTEACHING REVIEW CONFERENCE INTERACTION ANALYSIS SYSTEM

Categories for the content dimension

Category 1: Structuring (STR)
All administrative and procedural comments e.g. directing and leading procedures which set the context for discussion by launching or halting interaction.

Category 2: Teaching skill under focus (TS)
The particular teaching skill practised in the microlesson under review; talk about the skill, its performance in the microlesson, and future application.

Category 3: Other teaching skills and teacher behaviours (OS)
All references to other specific teaching skills, global skills and other aspects of the student teacher's behaviour in the microlesson.

Category 4: Lesson organisation and structure (LS)
All considerations of the microlesson structure as planned and as performed e.g. discussion of the lesson objectives, the less on plan etc.

Category 5: Lesson subject content (LC)
The subject matter of the lesson, its sequencing, accuracy, suitability, etc.

Category 6: Lesson materials and aids (LM)
Discussion of aids, apparatus etc. used in the microlesson.

Category 7: Pupils' characteristics and behaviour (P)
All talk about the microclass and individuals within it. Includes more general references to pupil's behaviour e.g. sociological and psychological perspectives.

Category 8: Discipline and control (D)
Classroom management, discussion of good examples or problems observed in microlesson.

Category 9: Next week's lesson (NWL)
Discussion about less on plan, content, materials and aids for the next lesson.

Category 10: Next week's skills (NWS)
Discussion about the teaching skill which is to be practised next week. Plans for other skills which need to be further practised next week.

Category 11: The microteaching system (SYS)
The operation and organisation of the microteaching system e.g. the objectives of microteaching and the review conference, the length of microlessons, the timetable etc.

Category 12: Other talk (OT)
All other talk. Include non-task orientated remarks and social exchanges.

Category 13: Silence (SIL)
Each separate pause estimated to be longer than three seconds. No attempt is made to estimate the length of silences.

Category 14: Non-codable (X)
Inaudible remarks and situations of confusion e.g. two people talking at once.

Figure 2b
THE MICROTEACHING REVIEW CONFERENCE INTERACTION ANALYSIS SYSTEM

Categories for the Problem Solving and Verbal Interaction Dimension

Category 1: Tutor gives information (TGI)
All supervisor talk which provides information e.g. factual comments, instructions, structuring procedures and directions, explanations of a teaching skill, etc.

Category 2: Tutor gives analysis (TGA)
Supervision statements which are classed as opinion, analysis or evaluation. Statements of praise or criticism made by the supervisor.

Category 3: Tutor gives suggestion (TGS)
Supervisor statements which give advice or offer suggestions.

Category 4: Tutor asks for information (TQI)
Tutor questions asking for a factual or descriptive reply.

Category 5: Tutor asks for analysis (TQA)
Tutor questions which ask the student (s) to give an opinion, judgement, analysis or evaluation.

Category 6: Tutor asks for suggestions (TQS)
Tutor questions which ask the students to offer advice, suggest improvement, modification, different approach etc.

Category 7: Tutor accepts, responds (TA)
Supervisor utterances which respond to student talk by accepting or using the student's ideas e.g. repeating, rephrasing, confirming, extending the student's comments. As the supervisor brings his own ideas into use the categorisation switches to category 1, 2 or 3.

Category 8: Tutor other talk (TOT)
All other supervisor talk e.g. comments which are not task orientated, comments which help build or disrupt the supervisory relationship, social remarks, etc.

Category 9: Student teacher no. 1 talks (ST1)
Students are numbered from left to right as seen on the videorecording of the review conference. This category is used for all distinguishable verbal utterances by the left-most student of the group.

Category 10: Student teacher no. 2 talks (ST2)

Category 11: Student teacher no. 3 talks (ST3)

Category 12: Student teacher no. 4 talks (ST4)

Category 13: Silence (SIL)
As in the content dimension

Category 14: Non-codable (X)
As in the content dimension.

As Fig. 2 indicates the system now consists of two dimensions, the Content and the combined Problem Solving - Verbal Interaction dimensions. Various orders of categories were experimented with to ease coding, increase reliability and, in the case of the second dimension, to facilitate interpretations of matrices of results by the method developed by Flanders. These points are discussed in subsequent sections.

Stage 3: Time, Thought and Reliability.

Coding of categories may be at fixed intervals such as every third second (Flanders, 1960; Blumberg, 1970) or at the end of every simple thought unit (Bales, 1950; Weller, 1971). A thought unit is regarded as the simplest meaningful statement which may be coded. Examples are simple sentences, or one word responses. Complex sentences involve two or more thought units. At first glance fixed interval coding appears easier and more reliable. A series of trials of intra-observer reliability yielded Scott (1955) coefficients of O.84, O.81 and O.73 for fixed interval and O.87, O.82 and O.83 for thought unit coding. Perhaps more important the 3 second interval at times encompassed several interactions by different speakers whereas using the thought unit enabled one to code all the interactions. Hence thought unit coding was adopted and more detailed measures of its reliability taken.

Samples from six conferences were selected for this purpose and three separate trials were conducted on each conference. For the content dimension six of these yielded results above O.85, nine approached this level and three were below it. The main difficulty was the boundaries between 2 and 4 and 3 and 4 (Fig. 2a). Therefore another coding rule was introduced to alleviate this difficulty!

The reliability trials of the Verbal Interaction system (Fig. 2b) yielded sixteen Scott coefficients above O.85, one approaching this critical value and one below it. The results for this dimension are considered highly satisfactory.

Stage 4: Handling the data.

The strings of **numbers** obtained from the coding observations may be analysed using a computer programme devised by B. McGarvey. The programme for the Content dimension analysis gives the number of tallies in each category and expresses them as a percentage of the total. From these it is possible to extract the percentage of the discussion concerned with teaching skills (Categories 2 and 3); the specific skill/general skill ratio (Category 2/Category 3); and the distribution of talk over other topics.

Table 1 sets out examples of the Content analysis of three microteaching conferences. The greater part of the conferences was devoted to skills (Categories 2 and 3) but the specific/general skill ratios were O.59, O.43 and O.O8 respectively. The supervisors were supposed to focus primarily upon the skill under review, it is clear that one of them focused upon every other skill but the one under review. The same group, Group C, also did not discuss the development of lesson content and skills for the next lesson.

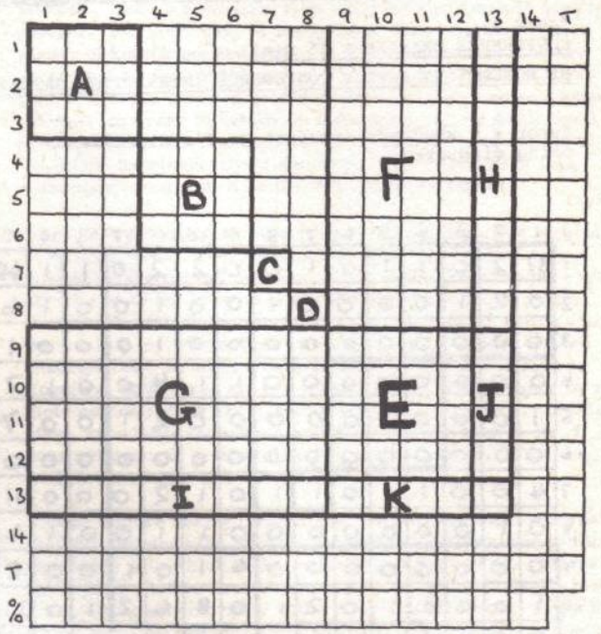
TABLE 1 = CONTENT ANALYSIS (Per cent)

	-I	2	3	4	5	6	7	8	9	IO	II	I2	I3	I4
A	4.9	24.3	41.3	I6.0	0.0	0.0	0.0	0.0	5.4	5.3	2.4	0.5	-	0.0
B	3.0	I8.2	42.0	I4.4	0.0	3.0	3.0	0.0	I2.9	.8	0.0	I.5	-	I.5
C	I.9	5.6	73.0	5.2	I0.0	0	0.0	0.0	0.0	0.0	I.5	I.2	-	0.4

(At the stage of development from which these examples are taken silence was not being recorded).

The programme for the Verbal Interaction dimension gives the number and percentage of thought units and a matrix of the two step chains of interaction (Flanders, 1970). Fig. 3 sets out the main areas of the matrix. Areas A, B, C, D and E are described as steady state areas. They indicate extended use of the same category. A gives the supervisor talk which may be described as telling, B with questioning, C with responding by accepting student's contributions and D with other forms of supervisor talk. Area E indicates the extent of student-student discussion in which the supervisor is not

FIGURE 3:
MICROTEACHING REVIEW CONFERENCE INTERACTION MATRIX



°/o TUTOR TALK _ TUTOR INFORMATION RATIO
°/o STUDENT TALK TUTOR ANALYSIS RATIO
TUTOR-STUDENT TALK RATIO TUTOR SUGGESTION RATIO
TUTOR INDIRECT-DIRECT RATIO

involved. The columns of area F give a comparison of how students react to the supervisor and the rows of G how the supervisor reacts to each of the students. H gives the nature of the supervisor's behaviour which is followed by silence and I how the supervisor reacts to silence. Area J indicates which student's remarks are followed by silence and K shows which student ends the silence. The programme also supplies various indices and Fig. 4 shows how they are compiled. In addition it is possible to pick out behaviour patterns from the matrix which are characteristic of various supervisors.

Fig. 4 SOME VERBAL INTERACTION INDICES

percent tutor talk	$= \frac{\text{sum of columns I-8 x IOO}}{\text{sum of columns I-I4}}$
percent student talk	$= \frac{\text{sum of columns 9-I2 x IOO}}{\text{sum of columns I-I4}}$
tutor-student talk ratio	$= \frac{\text{sum of columns I-8}}{\text{sum of columns 9-I2}}$
tutor information ratio	$= \frac{\text{sum of columns I+4}}{\text{sum of columns I-8}}$
tutor analysis ratio	$= \frac{\text{sum of columns 2+5}}{\text{sum of columns I-8}}$
tutor suggestion ratio	$= \frac{\text{sum of columns 3+6}}{\text{sum of columns I-8}}$
tutor indirect-direct ratio	$= \frac{\text{sum of columns 4-8}}{\text{sum of columns I-3}}$

Tables 2, 3 and 4 set out the matrix and indices of three conferences. They are presented here as illustrations and not as representative samples of the analyses of microteaching supervision.

Conference Example A (Table 2)

This tutor talked for 52% of the total thought units, and was mostly concerned with information giving and questioning.

TABLE 2:
CONFERENCE EXAMPLE A (5 min)
MICROTEACHING REVIEW CONFERENCE INTERACTION MATRIX

Tutor + 4 students; student no.3's microlesson being discussed.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	T
1	37	2	0	1	1	0	1	0	0	2	2	0	1	1	48
2	0	2	1	0	0	0	0	1	0	0	1	0	0	1	6
3	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1
4	0	0	0	0	0	0	0	0	1	1	4	0	0	1	7
5	1	0	0	0	3	0	0	0	0	0	2	1	0	0	7
6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	4	0	0	1	1	0	1	1	0	1	2	0	0	0	11
8	0	1	0	0	0	0	0	0	0	1	1	0	0	1	4
9	0	0	0	0	0	0	0	0	6	1	0	1	0	0	8
10	1	0	0	1	1	0	2	1	0	8	6	2	1	0	23
11	3	0	0	3	0	0	6	0	1	5	9	0	0	1	28
12	1	0	0	0	0	0	1	0	0	3	0	7	0	0	12
13	1	0	0	0	0	0	0	1	0	0	0	0	1	0	3
14	0	1	0	1	1	0	0	0	0	1	0	1	0	0	5
T	48	6	1	7	7	0	11	4	8	23	28	12	3	5	163
%	29.5	3.7	0.6	4.3	4.3	0.0	6.8	2.5	4.9	14.1	17.2	7.4	1.8	3.1	

% TUTOR TALK	51.53	TUTOR INFORMATION RATIO	0.66
% STUDENT TALK	43.56	TUTOR ANALYSIS RATIO	0.16
TUTOR-STUDENT TALK RATIO	1.18	TUTOR SUGGESTION RATIO	0.01

TUTOR INDIRECT-DIRECT RATIO 0.53

The tutor analysis ratio was low and the tutor suggestion ratio was very low, so this sample of conference behaviour does not have a problem solving perspective. The indirect-direct ratio of 0.53 shows the tutor's style to be rather more direct than indirect. All four students contributed to the discussion with the student whose microlesson was being discussed contributing most. The pattern of interaction was mostly student-tutor-student with some student-student exchanges, particularly between students no. 2 and no. 3. Area F shows that the tutor directed his remarks almost exclusively to the micro-teaching student. There was an appreciable amount of non-codable interaction in this conference.

Conference Example B (Table 3)

TABLE 3:
CONFERENCE EXAMPLE B (14 min)
MICROTEACHING REVIEW CONFERENCE INTERACTION MATRIX

Tutor + 4 students; student no.4's microlesson being being discussed.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	T
1	20	1	1	1	0	0	0	0	0	1	0	4	1	0	29
2	3	19	1	1	5	0	0	0	0	1	0	4	0	0	34
3	1	1	9	0	0	0	0	0	0	0	0	0	0	0	11
4	0	1	0	9	0	0	0	0	2	3	1	10	1	0	27
5	0	0	0	1	8	0	0	2	1	1	0	6	2	1	22
6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	1	2	0	5	1	0	13	0	1	1	1	7	3	0	35
8	0	0	0	0	1	0	0	0	0	0	0	1	0	0	2
9	0	0	0	1	0	0	2	0	4	2	0	0	0	0	9
10	0	1	0	2	1	0	5	0	1	22	0	2	0	0	34
11	0	0	0	1	0	0	2	0	0	0	1	0	0	0	4
12	4	9	0	4	4	0	12	0	0	3	0	25	2	0	63
13	0	0	0	2	1	0	1	0	0	0	1	4	0	0	10
14	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0
T	29	34	11	27	22	0	35	2	9	34	4	63	10	2	281
%	10.3	12.1	3.9	9.6	7.8	0.0	12.5	0.7	3.2	12.1	1.4	22.4	3.6	0.7	

% TUTOR TALK	56.94	TUTOR INFORMATION RATIO	0.35
% STUDENT TALK	39.15	TUTOR ANALYSIS RATIO	0.35
TUTOR-STUDENT TALK RATIO	1.46	TUTOR SUGGESTION RATIO	0.07

TUTOR INDIRECT-DIRECT RATIO 1.16

This tutor talked for 57% of the total thought units and had a fairly indirect approach (indirect-direct ratio = 1.16). This was achieved by having a large measure of tutor encouragement and acceptance (category 7; 12.5%) and by having 3.6% of silences which were mostly ended by students offering comments. The tutor was concerned with analysis as well as with information, but did not give or ask for many suggestions. There was very little student-student interaction and the interaction pattern was to and from the supervisor. The microteaching student spoke more than the other students and one other student took a major part in the discussions. The other two students made very small contributions and only participated when invited to do so by the tutor.

Conference Example C (Table 4)

TABLE 4:
CONFERENCE EXAMPLE C (8.5 min)
MICROTEACHING REVIEW CONFERENCE INTERACTION MATRIX

Tutor + 3 students; student no. I's microlesson being discussed.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	T
1	78	0	3	2	2	0	0	0	3	0	0	0	2	0	90
2	0	11	0	0	1	0	0	0	0	0	0	0	0	0	12
3	1	0	16	0	0	0	0	1	2	0	0	0	0	0	20
4	0	0	0	1	0	0	0	0	2	0	7	0	0	0	10
5	1	0	0	0	4	0	0	0	3	2	0	0	3	0	13
6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	4	0	1	1	0	0	0	0	0	1	0	0	0	0	7
8	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1
9	4	0	0	0	2	0	4	0	10	1	0	0	0	0	21
10	0	0	0	0	2	0	1	0	1	2	0	0	0	0	6
11	0	0	0	5	0	0	2	0	0	0	5	0	0	0	12
12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	1	1	0	1	2	0	0	0	0	0	0	0	1	0	6
14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
T	90	12	20	10	13	0	7	1	21	6	12	0	6	0	198
%	45.5	6.1	10.1	5.1	6.6	0.0	3.5	0.5	10.6	3.0	6.1	0.0	3.0	0.0	

% TUTOR TALK	77.27	TUTOR INFORMATION RATIO	0.65
% STUDENT TALK	19.70	TUTOR ANALYSIS RATIO	0.16
TUTOR-STUDENT TALK RATIO	3.92	TUTOR SUGGESTION RATIO	0.13

TUTOR INDIRECT-DIRECT RATIO 0.35

During this sample of discussion the supervisor was very direct in style (indirect-direct ratio = 0.25) and did 77% of the talking. He was concerned mostly with information giving, and also gave some analysis and suggestion, as shown by the heavy concentration of tallies in area A of the matrix. Area E and F show that the role of the students was simply to respond to the tutor. All periods of silence in this sample followed after tutor talk and were terminated by the tutor speaking again. There were two types of silence; those tallied in cell 1, 13 which were hesitations in the tutor's flow of speech, and those tallied in cell 5, 13 which were silences following the tutor's analytical questions which areas I and K show remained unanswered by the students.

These three examples suggest that the system may yield a wealth of interesting and revealing information on the micro-teaching supervisory conference. Different patterns of behaviour can be described and analysed, especially in relation to the verbal behaviours of the conference supervisor.

Summary

This paper has described the development of an instrument for exploring the supervisory process in microteaching. The instrument has relatively high observer reliability. It yields analyses of both the Content of supervision and the Verbal interaction processes involved. Its use of a matrix method permits sensitive analyses of the supervisory conference. Some supervisory conferences have been analysed using the instrument and others will be analysed during 1975-6. The next stage will consist of exploring the processes of supervision in conjunction with changes in the teaching skills of the students being supervised.

BIBLIOGRAPHY

BALES, R.F., (1950), 'Interaction process analysis: a method for the study of small groups," Addison-Wesley

BLUMBERG, A., (1968). "Supervisor behaviour and interpersonal relationships, "Educl. Admin. Quart., 4, 34-45.

BLUMBERG, A., (1970), "Supervisor-teacher relationships: a look at the supervisory conference, " Administrators' Notebook, 19, 1-4.

BROWN, G.A. (1975), "Microteaching: Programme of Teaching Skills, "Methuen.

BROWN, R.V. and HOFFMAN, M.J.S., (1966), "A promissory model for analysing and describing verbal interaction between college supervisors and student teachers during supervisory conferences, "Ed. D. dissertations, University Microfilms 67-2787, Ann Arbour, Michigan.

DARR, A.D., (1973), "Predictors of supervisor-teacher conference interaction, "ERIC ED 075 912

DUSSAULT, G., (1970), "A theory of supervision in teacher education," Teachers' College Press.

FLANDERS, N., (1960), 'Interaction analysis: a technique for quantifying teacher influence, "Paper presented to the A.E.R.A. meeting.

FLANDERS, N., (1970) "Analysing Teaching Behaviour, " Addison-Wesley.

HEIDELBACH, R.A., (1968), "A tentative model for analysing and describing supervisory teaching, " Classroom Interaction Newsletter, 3, 15-20.

McCROSKEY, J.C. and WRIGHT, D.W., (1971), "The development of an instrument for measuring interaction behaviour in small groups, " Speech Monograph, 38, 335-340.

MICHALAK, D.A., (1968), "An exploratory study of supervisor verbal behaviour in a student teacher conference, "Ed. D. dissertation, University Microfilms 69-14678, Ann Arbour, Michigan.

PARRY, G., and GIBBS, I., (1974), "A bibliography of supervision, "Prog. Learning, 11, 97-111.

SCOTT, W.A., (1955) 'Reliability of content analysis: the case of nominal coding, "Public Opinion Quarterly, 19, 321-325.

WELLER, R.H., (1971), "Verbal Communication in instructional supervision, "Teachers' College Press, Columbia University.

AN ANALYSIS OF THE COSTS OF MICROFILMING

Ken Kennedy
Social Development Department
Peterborough Development Corporation

INTRODUCTION

1. Since microteaching began in this country in 1969 there has been considerable enthusiasm shown for it among teacher educators. Some of the reasons for this maybe:

(a) that it can be deliberately planned to utilise the conclusions of psychological research into the acquisition of new patterns of behaviours;

(b) that the assumptions which it embodies, the procedures it involves and the goals towards which it is directed, have generally (if sometimes crudely) been made explicit;

(c) that systematic research into it are much more possible than other innovations in teacher education thus making possible its own refinement;

(d) that it provides a context in which it is possible for the theoretical study of teaching and learning to be related directly to the practice of teaching; an outstanding weakness of teacher education programmes in recent years.

2. This is of course not evidence in itself but there is evidence that microteaching exercises can lead to substantial changes in specific aspects of students' teaching behaviour. In spite of this few institutions in Britain have incorporated microteaching into the general programme of professional education. Why is this so?

3. The reasons may be that microteaching requires considerable advance preparation, even restructuring of courses, it requires consideration of timetabling, accommodation, resources and staffing and it is believed that it is prohibitively expensive. Such a view might be based on the assumptions that microteaching necessarily involves video-tape feedback, that, if television equipment is used, it must be as complex as that used at the University of Stirling and that it involves a commitment to a one-to-one relationship between student and supervisor, thereby again raising costs far beyond what colleges would expe ct to spend on college based activities for students.

4. Are these fears justified? Is microteaching significantly more expensive than other techniques? Because the precise objectives of other techniques are difficult to compare with microteaching objectives it is not possible to answer this question as put. What one can ask however is whether the extra benefits of using microteaching justifies any extra cost. To answer this in economic terms would entail longitudinal data on:

(a) the nature, extent and permanance of teachers' behavioural changes;

(b) the nature, extent and permanance of the effect of these differences of teachers' behaviours on pupils knowledge, skills, and attitudes;

(c) the economic value of such differential pupil gains.

5. In the absence of such information we are thrown back to reconsider our basic question and to investigate not broad teacher educational programmes but differing combinations of the factors which go to make up microteaching procedures. Because microteaching is not unitary it is likely that while some of the combinations of contributing components are expensive less complex ones could be quite inexpensive, and since cost information alone tells us virtually nothing it is necessary also to investigate the relative effectiveness of techniques. That is we require to identify those types of programmes which are most effective at any given level of expenditure and to discover if these costs can be reduced without sacrificing effectiveness.

6. This presents us with the problem of effectiveness and we have attempted to overcome it more by what appeared to be feasible than by what could be theoretically desirable. Assessment then of the relative effectiveness of different microteaching "treatments" have to be based on the judgements of a sample of experts - people whom we had reason to believe were both experienced in the use of some of these treatments and were thoroughly conversant with the current research and literature on the subject.

7. As to the criteria by which such judgements could be made we felt that it would not be possible to specify them because of the large number of possibilities there are for people to choose from, no way for us to know what exactly they might be and no basis for deciding upon the relative weights of each criterion. The judges were therefore asked to make their assessments on whatever criterion of effectiveness, or compound of criteria, each of them believed to be generally most relevant for microteaching in pre-service teacher training.

8. Our conclusions then about cost-effectiveness will depend on implicit consensus among our "experts" as to what criteria are important. We will simply be concluding that a particular treatment is the most effective in achieving whatever it is that microteaching is useful for.

9. It was clearly impossible however to ask for effectiveness ratings of all treatments, since they number in the thousands, but since there are a number of identifiable components which combine to make up treatments we proceeded to investigate the effectiveness ratings of these components (C.50) and their cost estimates. An assessment of the costs and effectiveness for any treatment could then be obtained by totalling the cost and effectiveness measures for the components which contribute to the treatment.

10. This list of components had to be comprehensive enough to take account of all possible microteaching treatments; components had to be defined in terms which could be costed and which the judges would feel themselves able to rate. In actuality the final list was arrived at by successive approximations.

11. Whilst the definition of the components was carried out in conjunction with the development of the rating schedule this paper refers mainly to the economic aspects of the analysis.

ASCERTAINING COSTS

12. It is apparent from the above then that the next step was to analyse our pedogogical definitions of potential components of a microteaching programme in a way which enables us to say precisely what in physical terms went to make up a component eg. component 45 "supervisor trained to use systematic observation schedule for skill (eg. 2 hours per skill of seminar workshop)". Even the final parenthesis of that component gives you a clue to how we attempted to define components in ways which could be costed. We concluded that this component consisted of lecturers' time (preparation and presentation); use of models; trainees' time, practice equipment and accommodation. Then, as in the parenthesis, we made an estimate of the length of time required for supervisors to be trained and acquire such skills, as would be useful to them in commenting on students' work, based upon our own experience at Stirling and on such evidence as was available in the literature on training of supervisors for microteaching [2].

13. A similar analysis was carried out with each component at every stage; each contributing item was subjected to scrutiny to ascertain if cost data had already been found for it or if it could be found.

14. It was of course necessary to find a lowest common denominator to all costs which was relevant to our effectiveness ratings. A choice lay among, for example, cost per student, cost per programme, cost per hour or cost per lesson. Since however we were requesting respondents in our questionnaire on factors influencing the effectiveness of microteaching to rate the importance of the contribution each

component would make per microteaching lesson it was appropriate that, no matter how difficult to conceive of, in some instances, or how difficult in practice to ascertain such figures we were constrained to choose a per lesson cost basis. Such difficulties as this poses will be described as appropriate below.

15. As might well be imagined costs per microteaching lesson of all the innumerable items which go to make up a programme were not lying around waiting for us to come along. As an initial step then to discover such data we distinguished among four types of inputs.

A. Fixed capital assets eg. buildings.

B. Moveable capital assets eg. equipment.

C. Recurrent inputs eg. staff and transportation.

D. Non institutional and non-economically identifiable inputs.

Of these, in any institution, we were only concerned with those items which were potentially new inputs into existing teacher education courses, ie given a particular course what would be the cost of adding on each or any of the 50 components. Let us look therefore briefly at each of the major sub-divisions above.

16. A. Fixed Capital Assets have to be included, in spite of the fact that they probably have been already provided, because additions to programmes will in most cases require a variety of resources not least of which is likely to be accommodation. To reduce the gross cost of a building in which any component of microteaching is conducted it was necessary to firstly multiply the initial cost of the building by an appropriate sinking fund factor (which reflects a rate of interest on money borrowed over a large number of years) to arrive at the annual capital value. This if further divided by the total area of the building to discover the cost/sq ft/annum.

$$\text{ie cost/sq ft/annum} = \frac{\text{initial cost x sinking fund factor}}{\text{area of building}}$$

and the cost/hour of space is calculated by the following formula.

$$a = \frac{S(c + b + d + e + f)}{H} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (i)$$

where
- a = cost/hour of space
- S = area of space required
- H = number of hours per year space could be used
- c = capital cost/sq ft/annum
- b = portering cost/sq ft/annum
- d = cleaning cost/sq ft/annum
- e = maintenance cost/sq ft/annum
- f = power cost/sq ft/annum (including heat)

This figure, a, above is used as an input into several components as will be shown later.

17. B. Moveable Capital Assets. Where equipment is concerned it is not sufficient just to know its initial cost but also over what period its cost can be written off (ie how long it will last and how much maintenance it will require during that time). From the initial cost of the equipment we calculated the cost per hour by the following formula.

$$c = \frac{x}{y\,t} + \frac{z\,k_1}{(y-z)} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (ii)$$

where
- c = cost/hour
- x = purchase price
- y = maximum number of hours in a working year item could be used
- t = number of years of expected life
- z = estimated number of hours of maintenance/year
- k_1 = estimated cost of Grade 5 CCTV technician
- $(y-z)$ = net number of hours/year the equipment can be used after maintenance

18. C. Recurrent Inputs. These depend upon the scale of the operation, are more diverse and thus have greater variability than previously discussed inputs. Academies' annual salaries were divided by 2254[3]. Others were assumed to work a 1725 hour year.[4]

Minor items of equipment are supplementary to moveable capital assets or are spares; they are assumed to have a life span of one year.

19. Where real pupils were used pupil transportation costs per day were divided by the number of microteaching lessons conducted in the institution using these pupils, to arrive directly at a "per lesson" cost estimate.

20. Materials used up by students or staff in preparing for lessons also had to be estimated directly on a per lesson basis. Expenses and subsistence to staff and students on regular teaching practice usually accounts for a huge part of a college's budget and would do so if microteaching were conducted in schools. If not however the item is likely to be confined to the administration and setting up of the programme - especially if real pupils are used.

21. D. Non-Institutional and Non-Economically Identifiable Costs.

One further category of costs remains to be discussed - those not incurred by institutions - but we will only do so briefly for so far as this particular study is concerned they are not of great importance. One could include, for example, the opportunity cost of students' time since microteaching is assumed to be an additional commitment they could presumably be doing something else. This cost is however borne by society and not the institution for it is society at large which supports students and loses their inputs on the labour market while they study.

22. A cost nearer to home, however, which every teacher educator is aware of, is that incurred by pupils, teachers and headteachers in schools who suffer, at the least, inconvenience by being persuaded to allow pupils to be used as guinea pigs whether in regular teaching practice or in microteaching. It might be then that schools suffer disruption by sending out their pupils to college and consequent upon this is the ill-feeling engendered between the school and the college leading to inefficiencies in future dealings with the school. It is obviously very difficult to quantify such costs and even to adequately identify them but similarly on the benefit side one ought also to be aware of potential spin-off for pupils and also for the college from goodwill engendered by liaison with schools.

COSTING COMPONENTS

23. Parameters. We have not attempted to specify a range of treatments but to present a large number of components which could be used in a microteaching programme and to lay out the items which contribute to any component. Inevitably, to specify the cost of the component in 'per lesson' terms involves parameters which will differentially affect inputs according to their place in a component. These parameters will obviously alter according to the peculiar characteristics of any one institution. Before we explain the components let me identify the parameters which are important.

1.	Number of skills practised in the programme by each student	S
2.	Number of lessons practised per skill	N
3.	Number of specialisations catered for in a skill	X_1
4.	Number of students practising a skill per programme run (term, year or semester)	M
5.	Number of tutors available for observation (or training)	T
6.	Number of programme runs	W
7.	Number of specialisations for models	X_2

24. In our enquiries we found these parameters to range widely as below, which gives an indication of how different costs could be for any other institution attempting to use the formulae for there is nothing sacrosanct about the ranges we found.

S varied between 1 and 6

N varied between 1 and 2

X_1 varied between 1 and 2

M varied between 25 and 350

T varied between 3 and 35

W varied between 1 and 3

X_2 varied between 1 and 4

25. Optional Extras. Because of the extreme complexity of most of the components and because they could be taken to be compiled of a number of inputs which could be more or less expensive depending on how one interpreted the resource requirements of a component we concentrated on the fewest, simplest and cheapest resources which could possibly combine to make up a component. That is, while most of the seven colleges used more resources per component than was absolutely required we have costed components on an addition of the bar minimum of elements necessary. There is therefore a gap between the theoretical minimum and that which was currently employed and where we were aware of such a difference and were able to state precisely the nature of the improvements over the basics we have termed these the 'optional extra' strategies of achieving a component.

26. If, in using our formulae then, one believes that our simplest system is not what one wants then we have attempted to suggest the improvements one could make - usually at greater expense, without utilising techniques or resources from more sophisticated components (those listed later in the questionnaire) and usually with no guarantee of improved effectiveness. For example, if one brings real pupils into the college, then one incurs transportation, administration, supervision, escorting and accommodation costs as a minimum. The "optional extras" we found some colleges to be employing were: materials for the pupil to use while waiting for their lessons, food and drink provided (especially if staying the whole day), more spacious accommodation and using the same pupils with the same student at each micro-teaching session (thus raising administrative costs, decreasing flexibility and increasing the number of unfilled microteaching studio 'spots' as pupil student matchings are unable to be made).

27. Component elements. As indicated in paragraph 12, above, components were analysed. This involved discussion with all those who took part in microteaching at Stirling - lecturers, administrators, office staff and technicians - to ascertain exactly what resources were required and for how long. I say Stirling but this was only in the initial stages as we developed the general statements for each component, thereafter all the questions about procedures, about equipment, about who carried out jobs etc were investigated at several colleges. To illustrate how this information was used let me use one of the components; No. 7 "Modelling of Skills through Audio Tapes (assuming no video-tape models)".

28. Firstly we know roughly that a lecturer has to think about the behaviour he wishes to illustrate, record several lessons, listen to the replays and pick out the sections he requires, prepare his lecture and present it. Experience suggests that three ten minute models have been commonly used. We assume the lecturer would require to do this. Secondly we know he will require an audio taperecorder for recording, selecting and presenting his lecture and thirdly this machine will require tape for these activities and while it does last a long time it does deteriorate with use and eventually has to be written off.

29. The next step is to determine how much a lecturer costs per hour, how much a taperecorder costs per hour and how much a tape costs per pass through the machine. We can discover the first by dividing the total cost of a lecturer's average salary, superannuation and insurance (i.e. his cost to the institution) by the number of hours per year on average a lecturer works[5]. We can calculate from formula (ii) above the cost/hour of the machine and we can make an estimate for the number of times a tape can pass through a machine.

30. That deals with 'what resources?', now how long are they utilised for? Estimates from different institutions vary considerably but it is the only way apart from observing a lecturer actually thinking, recording, checking etc. We did however use a formula, developed at Stirling for estimating staffing workloads, to place a time on lecture preparation and presentation.[6]

31. These elements and their relationship can be expressed in the following general statement.

$$C_7 = \frac{k(Tw + Tr + Tc + L + \frac{1}{2}Lp) + e(Tr+Tc+L) + P(Pr+Pc+Pl)}{M\,N\,W} \quad \text{(iii)}$$

where

C_7 = cost of component No 7
k = Salary of a lecturer/hour
Tw = Time to write out behaviour to be illustrated
Tr = Time to record lessons
Tc = Time to check and select from tapes
L = Time to present lecture
Lp = Time to prepare ordinary lecture
e = Cost/hour of audio taperecorder
P = Cost/pass of audio tape
Pr = No of passes required for recording lessons
Pc = No of passes required for selection
Pl = No of passes required for lecture presentation
M = No of students practising the skill per programme run
N = No of lesson practice per skill
W = No of programme runs.

I make no apology for not particularising this with figures at this stage for it is the general which is important and furthermore the values of the parameters M, N and W have to be decided in a particular case.

32. From 7 the analysis proceeds to 8 which makes the general statement of the additional costs over 7 one could expect to incur, by using video tape instead of audio, and so on through the whole questionnaire; building up increasingly sophisticated components and determining at each what the new elements were over previous components to ascertain additional costs.

COST FINDINGS

33. Once all the general statements had been developed and costed for Stirling University, as a pilot run, we then collected similar cost data and running arrangements from six other colleges, viz:

Callander Park College of Education (Falkirk)
Jordanhill College of Education (Glasgow)
New University of Ulster (Coleraine)
Northern Ireland Polytechnic (Jordanstown)
Berkshire College of Education (Reading)
Eastbourne College of Education (Eastbourne)

which we knew to conduct microteaching either as a research project or better as an integral part of its teacher training programme; this in an effort to maximise the general usefulness of the formulae and the data and to minimise the idiosyncratic nature of Stirling University's microteaching set up.

34. It is not possible to lay out here the full matrix of costs for every item investigated at every institution, primarily because of space but also to avoid crude comparison being made between institutions on the basis of a few cost estimates whose contexts are not fully known and to avoid the particular instead of concentrating on the general. Suffice it is to say that the range was often large (see Appendix A) due probably to a variety of factors among which are the inherent differences in the educational systems of England, Scotland and Northern Ireland which lead to differences in methods of conducting teacher training and of course in salaries.

35. All the values (see Appendix A) were costed in £/hour as illustrated earlier then applied to the general statements of all the components and further reduced to cost per lesson depending on the values of parameters chosen.

36. I have attempted to confine myself to the economic implications of different components but it is necessary to say a final word about the effectiveness side of the research. As was indicated in the introduction we asked experts in microteaching to rate the additional effectiveness of a component over the previous component on a four point scale for, in spite of fact that we appreciate that further empirical research is required, we are also aware that policy decisions are rarely able to be made directly from research findings and that they depend largely on the insights of the practitioner.

37. The questionnaire was designed firstly by asking those in the Department of Education at Stirling who had been most involved with setting up of microteaching for those variables which could influence the effectiveness of microteaching and by discussing with them exactly what they meant. Once

refinements were made (in the light of discussion and with the constraint of being able to cost the component) the questionnaire was drawn up such that preparation components were listed first, followed by lesson components, then by technical feedback components and finally by inter-personal feedback components and also such that if component y depended on component x then component x appeared first.

38. A pilot run of the questionnaire was carried out on the fifteen members of the Department of Education of the University of Stirling, from which feedback further refinements were made to questions, order, presentation, etc.

39. The final selection of judges was made on the criteria that they be experienced in the use of microteaching and they should have a thorough knowledge of the research literature on microteaching. But while the cost analysis confined itself to the UK, the judges were selected from all over the world, especially the USA where the fuller use had been made of the techniques than in any other country. In all 61 replies were received from about 78 sent out and while there is no pretence at randomness or stratification we can claim that the selection is broadly based and that the consensus is authoritative for the mean number of components not rated was only 1.5. As an indication of why some of the components were not rated, two of the Australian respondents said that they questioned the 'skills' approach and one American that "These are research questions, aren't they?"

40. What we will now do is to combine these two sets of data in an effort to identify the most effective programmes within given cost constraints. The full analysis of the questionnaires and of the cost-effectiveness of components will be available shortly in our final report.

[1] This paper is derived from a research project conducted by Donald McIntyre and myself into the costs and effectiveness of microteaching during 1974 at the University of Stirling by courtesy of a grant from the Social Science Research Council.

[2] See for example: Griffiths, Roy - The Training of Microteaching Supervisors. Department of Education, University of Stirling 1974 mimeo.

[3] Committee of Vice-Chancellors and Principals of Universities of the UK - Report of Enquiry into use of Academic Staff Time p. 14 "education".

[4] Arrived at by multiplying 37.5 hours per week by 46 weeks.

[5] ibid p.14

[6] Cottrell T L, Staff Allocation in the University of Stirling Appendix "Method of Calculating Workload" Paper AA(72)2

APPENDIX A

COST OF ITEMS USED IN MICROTEACHING COMPONENTS 1972-73
(in £/hour unless otherwise stated)

Item	Range	Mean
SALARIES		
Academic	1.91 - 1.353	1.6356
Technician	1.53 - 1.0082	1.2057
Secretarial	.868 - .6377	.7435
Clerical	.6727 - .4568	.5693
Escorts	.8388 - .4342	.7365
Graduates	1.53 - .875	1.135
Supervisors	No range	1.00
ACCOMMODATION		
Lecture Theatre	.8901 - .6348	.7625
Laboratory	.2740 - .216	.246
Large Room	.1142 - .092	.1031
Medium Room	.0680 - .0571	.0625
Seminar or Small Room	.09 - .037	.0577
Studio	.56 - .1296	.2912
Microteaching Room	.1840 - .1142	.1426
Portering /sq ft	.1889 - .1100	.1495
Cleaning /sq ft	.2395 - .0542	.1469
Maintenance /sq ft	.2088 - .0302	.1195
Heating) Power /sq ft Lighting)	.2014 - .0932	.2002
EQUIPMENT		
½" TV Camera	.0475 - .0211	.0305
Monitor (sound)	.0371 - .0213	.0301
½" VTR	.05 - .0384	.0422
Tripod	.0044 - .0017	.0030
½" Adaptor Box	.0042 - .0007	.0025
Microphone	.0055 - .0014	.0038
½" Camera Cable	.0055 - .0004	.0030
Minimum Recording Equipment	.1191 - .0815	.1073
Minimum Replay Equipment	.0871 - .0633	.0739
½" Videotape/pass	.0188 - .0104	.0153
Audio tape/pass	.0024 - .0020	.0022
Audio Recorder	No range	.0205
Audio Pulser	No range	.0391
1" Videcan Tube	No range	.0025
Camera Lens (remote control)	.0363 - .0318	.0341
Camera Lens (zoom, manual)	.0291 - .011	.0204
Synchronising Pulse Generator	.04 - .0131	.0273
Dolly and Controls	.0300 - .0036	.0125
Pan and Tilt Head	.0326 - .014	.0235
½" VTR Editor	No range	.0507
1" Camera	.068 - .0255	.0515
Switch/Fader	.0587 - .0161	.0376
1" Camera Cable	No range	.0044
1" Tape/pass	.05 - .030	.0405
Plumbican Tube	.1233 - .0110	.1114
1" VTR	.173 - .0841	.1003
Monitor (picture only)	.0350 - .0207	.0267
Cut Box	No range	.001
Effects Generator	.055 - .025	.0400
Talk Back Equipment	.0037 - .0007	.0022
½" Editing Equipment	No range	.0808
1" Lens	No range	.0029

SOME ASPECTS OF EVALUATION

CHARACTERISTICS OF SUCCESSFUL STUDENTS ON A RESEARCH METHODS COURSE

John Bynner and Joan Whitehead.
Open University.

INTRODUCTION

Methods of Educational Enquiry is a third level Open University course designed to equip students with the appropriate knowledge and skills to evaluate educational research findings. The course comprises six blocks of correspondence material moving from Research Design through Data Collection, Data Analysis, Experimental Design to Evaluation and Assessment of Research. These blocks are accompanied by television and radio programmes, exploring in greater detail some of the concepts and techniques treated in the texts; and a project involving the collection of data from schoolchildren, which runs through the first three quarters of the course. This is written up as one of the students' tutor marked assignments. The course is assessed in two ways. Continuous assessment, consists of four tutor marked assignments (TMAs) comprising one on setting out the case for a piece of research of the student's own choosing, one on the design of research instruments for a specified research project, one on the student project, and finally one on evaluation of published research articles. There are also four computer marked assignments paralleling the tutor marked assignments. Terminal assessment consists of an examination in three parts, a computer marked part (one hour), a prepared section on three research articles that the students had received before the exam (one hour), and an essay on some topic in research methodology or philosophy (one hour). The students' final grade for the course is based on a conflation of grades achieved from the two forms of assessment (5O% exam, 5O% continuous). In this conflation, computer marked assignments carry one quarter of the weight of the tutor marked assignments.

Now that the course is into its third year it seemed opportune to examine the relationships between students' performance in these different forms of assessment and their final grade in the examination, to see what skills appeared to be at the highest premium for success in the course. In addition, it is clearly of interest to know, especially in the O.U. context, whether the previous educational experience that students bring with them before entry into the O.U. and during it, plays a part in determining how well they do in a particular course. Accordingly this analysis was carried out to answer the following main questions.

1. How does student performance on the different forms of assessment relate to the final grade they obtain and which assignment is the best predictor of final grade?

2. What skills are required for top performance on assignments that best predict final grade?

3. How does the students' previous educational experience both before joining the O.U. and since doing so, relate to the final grade they obtain on the course?

4. Does the work load of students during the year they are taking E341 affect their performance on the course.

METHOD

One of the principal difficulties in drawing conclusions from an analysis of the relationships between performances in Open University assignments is that the students have a choice as to which of the assignments they complete. In Methods of Educational Enquiry they were expected to complete three out of the four tutor marked assignments (including, in 1974, the one on the project, TMAO3) and three out of four computer marked assignments. This means that correlations between different forms of assignment need not necessarily be based on the same samples of students. And this seriously hinders any attempt to compare the correlations in any absolute sense, as it is impossible to tell whether choice of assignment is itself related to a distinct type of skill. In order to get over this problem the analysis was repeated a number of times on different samples of students to see whether there was any converging evidence that certain relationships between assignments were stable over all sampling situations. First of all two 1O% random samples of students (N=78, 75) were selected from the 1973 student body, stratified to be representative of all combinations of performance on continuous assessment and examination in the final conflation. In 1974, samples were again replicated this time taking the whole student population and dividing at random in two (N=212, 2O9). Finally, on the assumption that those students who completed the last tutor marked assignment, were most likely to have completed the previous ones, a final analysis was done of just those students (N=155). Identification of the skills being tested in each assignment was achieved in the case of the TMAs simply on the basis of the content of the assignment. In the case of the CMAs, item analysis was carried out on each of them, which enabled us to determine which items were the best discriminators in the test, and, correspondingly, the skills that were at a premium in it.

Student educational experience subsequent to and during the Open University was taken direct from the student record file. The following characteristics were examined: "highest level of educational qualification previously obtained before entering the Open University," "previous performance on Open University courses," and "current work load," in terms of the number of additional courses to Methods of Educational Enquiry that the student was currently doing.

Type of Assessment	1973	1974 Sample A	1974 Sample B	1974 completed TMA04
Exam CMA	.60(153)	.69(212)	.67(209)	.73(155)
Exam prepared	.66(153)	.53(211)	.62(209)	.57(155)
Exam essay	.69(153)	.58(211)	.57(209)	.64(155)
CMA 41(Research design)	.34(152)	.48(211)	.52(209)	.47(155)
CMA 42(Data collection)	.34(152)	.37(209)	.47(205)	.43(153)
CMA 43(Data analysis)	.25(148)	.37(208)	.34(205)	.40(153)
CMA 44(Design and analysis of Experiments)	.38(106)	.43(160)	.44(153)	.43(137)
TMA 01(Research design)	.48(149)	.41(209)	.50(202)	.39(148)
TMA 02(Data collection)	.44(145)	.52(192)	.64(187)	.60(126)
TMA 03(Data analysis)	.30(116)	.59(204)	.62(206)	.52(152)
TMA 04(Evaluation of research articles)	.50(79)	.63(85)	.58(81)	.60(154)

Note Numbers in brackets are sample sizes for which the correlations were computed.

RESULTS

1. Prediction of final grade. Table 1 shows the correlations between each of the forms of assessment and final grade for each of the samples for which the analysis was carried out (the two samples from 1973 were merged for the purposes of this table). It is notable that in 1973 the essay in the exam was the best predictor of final performance and in fact the three parts of the exam all proved to be better predictors than any of the other types of assessment. In view of the 5O% weighting of the exam in the conflation this result might be expected. Thus it is all the more surprising that in the two samples of 1974 students, it no longer held up. The best predictor continued to be one part of the exam the computer marked part, but following it came some of the TMAs, i.e., either TMAO2, TMAO3, TMAO4 were equally good predictors.

The CMAs also produced surprising results. Taking into account the fact that in the final grade the CMAs only carried a quarter of the weight of the TMAs, it is striking that the correlations for the four CMAs were not substantially lower than those for the TMAs.

The greater stability of the sample that did TMAO4 in 1974 is shown by the numbers in the brackets following each

correlation. It can be seen that for this more restricted sample the general picture does not change; it merely sharpens a little. Thus the exam CMA now rises clearly above all other forms of assessment as the best predictor of performance on the course, and following it come the exam essay and the two TMAs O2 and O4.

Table 2 shows the intercorrelation matrix for all the forms of assessment for the sample of students who completed TMAO4. The matrix is arranged in such a way as to demonstrate the extent of 'convergent and discriminant validity' (Campbell and Fiske 1959, Psychol. Bull 56, 81-106) of each assignment. It is notable that the relevant correlation between assignments across methods (in the diagonal in the body of table) are lower than the correlations between assignments with methods. This suggests that the TMAs and CMAs are tapping distinctly different types of skill at each stage, regardless of the content of the course to which they relate. It is notable that the exam CMA shows a higher level of correlation with all other forms of assessment than does any other form of assessment. In other words it seems to underpin the types of skills and knowledge crucial to success in the course.

TABLE 2

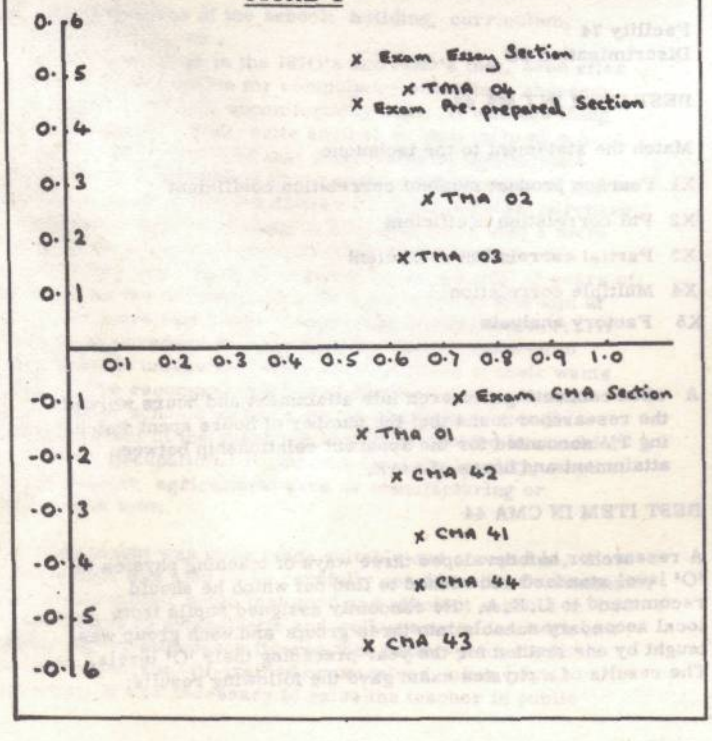
Correlations between different forms of assessm

Note Sample of student TMA04 in 1974.

N = 155

TMA 01										
TMA 02	0·22									
TMA 03	0·32	0·40								
TMA 04	0·25	0·49	0·44							
Exam: Pre-prepared Section	0·19	0·37	0·24	0·37						
Exam: Essay Section	0·21	0·34	0·28	0·49	0·41					
CMA 41	0·30	0·40	0·26	0·26	0·16	0·23				
CMA 42	0·30	0·31	0·29	0·26	0·28	0·18	0·36			
CMA 43	0·26	0·20	0·21	0·13	0·09	0·12	0·42	0·38		
CMA 44	0·36	0·28	0·32	0·27	0·22	0·10	0·53	0·39	0·42	
Exam: CMA Section	0·31	0·41	0·28	0·32	0·37	0·32	0·63	0·53	0·43	0·44

FIGURE 1

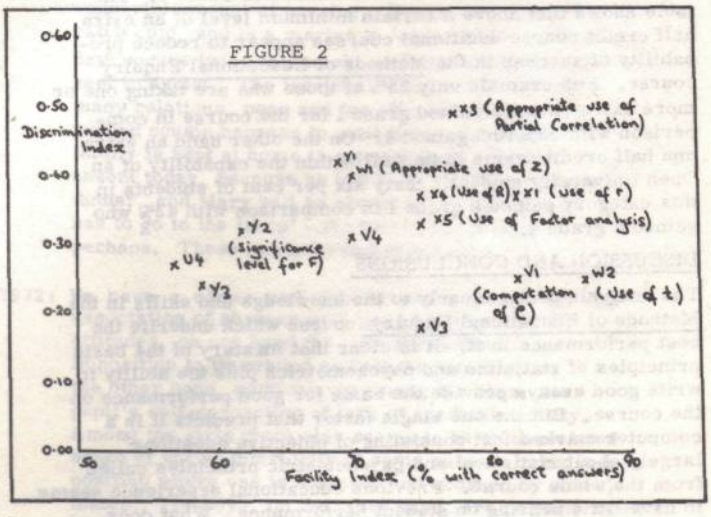


This conclusion is borne out by the results of a factor analysis of this correlation matrix (figure 1). Two principle components with latent roots greater than one accounting for some 50% of the total variance in the matrix were extracted. As can be seen from the location of the various assignments in the space defined by these two components, the exam CMA is the purest measure of the first general factor underlying all the tests. It is notable too that the assignments form two broad groups largely defined by method. (This is confirmed by a Varimax rotation of the first principal components). With the exception of TMAO1, which holds an anomalous position more closely aligned to the computer marked assignments, all other forms of assessment divide by whether they were tutor marked or computer marked. Thus even the examination splits clearly between the tutor marked part of it (the essays) and the computer marked part. Another possible interpretation that can be placed on the location of the assignments is that they separate computation skills from essay writing skills. CMA43 contains more computations than any other computer marked assignment and the exam essay is in some senses the most rigorous test in the course of the ability to write an essay.

2. SKILLS TAPPED BY BEST ASSIGNMENTS

It is clear from these results that the exam CMA appears to tap, more than any other, the skills essential for success in the Methods of Educational Enquiry course. Unfortunately at the time our work was being undertaken, it was not possible to carry out an item analysis of this CMA. On the other hand from an examination of the content of the other assignments, with which it was most strongly correlated, it is possible to infer the types of skills which were central to good performance on it. From table 2 it can be seen that CMAs 41, 43 and 44 all had correlations above .40 with the exam CMA, as did TMAO2 which was concerned with knowledge of the technical aspects of research instrument construction. Taking CMA43 to represent the rest, figure 2 displays the results of an item analysis of this CMA. On the vertical axis is the discrimination index for each item in the test; on the horizontal axis is the facility index. By an inspection of the items with the highest discrimination indices relative to their facilities, it is possible to see which are the key items in the test. It is noticable that those concerned with the concepts of correlation and the principles of testing for statistical significance are those which meet this criterion. Further inspection of the item analyses of all the other CMAs similarly reveals that the items doing best in each test are those concerned either with the understanding of statistical principles or psychometrics (reliability and validity). (The best items from the 4 CMAs are shown in the Appendix).

FIGURE 2



3. PREVIOUS EDUCATIONAL EXPERIENCE.

Table 3 relates students' final grade on the course to previous educational experience. First it is notable that previous educational qualifications seem to bear little relating to performance on the course. Certainly whether or not the student had university experience over and above teacher training college experience appeared to make no difference to his performance at all. It might be the case that the student with a background in science and maths would do better than others on a course demanding some facility with mathematical methods, at least at the level of algebra and computation. Again the table suggests no evidence for this. If anything the only tendency it

reveals is for those who had previously concentrated exclusively on education courses to do slightly better than the others. Finally on the question of performance in previous O.U. courses, a very strong relationship is revealed: the better the student had done either at foundation course level or at second level, the more likely they were to do well on the Methods of Educational Enquiry course. In other words one of the best predictors of performance on the course is performance in previous Open University courses.

Table 3 Final grades on course analysed by characteristics

Student Characteristic	Final Grade[1] 4 %	3 %	2 %	1 %
Highest educational qualification				
Above teachers certificate	20	16	24	18
Teacher certificate	80	67	66	74
No higher education	0	17	10	8
Previous subject studied				
Science/maths/technology	29	38	25	37
Social Science	8	4	3	6
Arts	5	6	6	5
Education	55	53	52	62
Performance in previous O.U. courses				
Distinction in foundation course	0	5	16	27
Merit or distinction on 2nd level courses	18	22	52	61
N	61	91	103	87

[1] For final grade students are grouped into four broad categories:

1 = Merit or distinction
2 = Good pass
3 = Fair pass
4 = Bare pass or fail

4. WORKLOAD. Finally the question of whether the work load the student undertakes at the Open University can affect his performance in a course is examined in Table 4. The table shows that above a certain minimum level of an extra half credit course additional courses appear to reduce probability of success in the Methods of Educational Enquiry course. For example only 25% of those who are taking one or more extra courses gained grade 1 for the course in comparison with 53% who gained 4. On the other hand an extra one half credit seems to be well within the capability of an Open University student. Sixty six per cent of students in this category gained a grade 1 in comparison with 43% who gained a grade 4.

DISCUSSION AND CONCLUSIONS

The analysis points clearly to the knowledge and skills in the Methods of Educational Enquiry course which underlie the best performance in it. It is clear that mastery of the basic principles of statistics and psychometrics plus the ability to write good essays provide the basis for good performance on the course. But the one single factor that predicts it is a computer marked test consisting of objective questions largely about statistical and psychometric principles culled from the whole course. Previous educational experience seems to have little bearing on student performance. What does predict it more than anything else is how well students have done on previous Open University courses and whether or not they are operating with the right workload. These findings suggest that mastery of the Open University learning system is one of the main determinants of success on any Open University course. This includes dropping out of it if the student feels he is unable to master the content (about 25% of the students did not complete the course). Those who decide the course is for them i.e. find themselves in tune with the material, and are able to master such central concepts in statistics as that of sampling error and probability distribution, are likely to do well on it.

Table 4 Final Grade analysed by work load

Work Load	Final Grade 4 %	3 %	2 %	1 %
E341 + another 1½ credit course	23	22	13	10
E341 + another 1 credit course	30	20	23	15
E341 + another ½ credit course	43	53	50	66
EE341 alone	5	6	14	9
N	61	91	103	87

APPENDIX
Best Items in CMAs

BEST ITEM IN CMA 41

Match the research situation to the sampling technique
X2 A researcher found that the sample number in one category of his design was too small for adequate statistical analysis, he therefore doubled its representation.

A random

B opportunity

C multistage

D stratified (proportionate)

E stratified (disproportionate)

F cluster

G systematic

Facility 82
Discrimination 56

BEST ITEM IN CMA 42

Which one of the five properties does a Guttman Scale have

A Judges are always used to select the items

B The scale has equal intervals

C The scale tends to embrace a narrow range of content

D The scale always has eleven points

Facility 74
Discrimination 30

BEST ITEM IN CMA 43

Match the statement to the technique

X1 Pearson product moment correlation coefficient

X2 Phi correlation coefficient

X3 Partial correlation coefficient

X4 Multiple correlation

X5 Factory analysis

A when conducting research into attainment and hours worked the researcher found that the number of hours spent watching TV accounted for the apparent relationship between attainment and hours of work.

BEST ITEM IN CMA 44

A researcher had developed three ways of teaching physics to 'O' level standard and wished to find out which he should recommend to L.E.A. He randomly assigned pupils from local secondary schools into three groups and each group was taught by one method for the year preceding their 'O' levels. The results of a physics exam gave the following results:

Method 1	Method 2	Method 3
3	2	7
4	6	7
5	6	7

W2 What is the most appropriate significance test for this design.

A t test for correlated sample means
B t test for independent groups
C One way analysis of variance
D Two-way analysis of variance
E None of the above

Facility 65
Discrimination 47.

SOME METHODOLOGICAL ASPECTS OF COLLABORATIVE RESEARCH

A. F. McPherson.
Edinburgh.

The programme of Collaborative Research that the SSRC is funding in the Centre for Educational Sociology at Edinburgh over the next four years will do research into educational decision-making and into achievement and choice at the secondary and tertiary levels. This needs saying because the remainder of this necessarily brief paper will confine itself to changes that we are making in the organisation of the research. These have some bearing on the political and methodological problems that occupied the first annual conference of this association.

EXPLORING THE CONSTRAINTS

We aim not only to do research but also to explore what happens when we try to extend the boundaries of a research team to encompass, in principle, an entire educational community. We propose to divest research knowledge of any credence it may receive solely from the exclusive control of the means of knowledge production by researchers in their several guises and then to press these resources to the service of the scattered, and hitherto underpowered, understandings which participants hold of the system and of their place in it. The progress and consequences of this attempted redistribution of control should help to clarify the relationships between the different bodies of knowledge that participants hold about the workings of education. It should also make more explicit the relationship between such knowledge and the exercise of control. William Taylor talks about the richness of the role-specific knowledge to be found in educational systems. But why is it role-specific? Must it remain so and what would happen if we tried to universalise it?

We are exploring, therefore, a variety of constraints on what can be known about any educational system. But this is more than an expensive exercise in public self-doubt. We believe that a relocation of the control of the machinery of research may enhance an understanding of educational phenomena that is impoverished by the division of labour between academic subjects, between institutions (schools, colleges, universities, examination boards, research organisations and government, both central and local) and between roles (teaching, administering, researching, examining and governing).

Through the Scottish Education Data Archive (SEDA) we have already set ourselves to weaken the argument that there are technically insurmountable obstacles to understanding educational practice. With one constraint removed, the nature of the remainder should be more evident to the evaluation component of the programme.

Several characteristics of educational systems commonly make it difficult to learn from them. How, firstly, can we hold cause and effect in the same field of vision when pupils are always passing through the settings that any one person can observe? Educational systems, secondly, tend to be scattered and private. How, then, can we begin to compare, contrast and generalise, to locate our local experience and understanding in the field of like situations, accumulating and discarding ideas in a systematic way? And How are we, thirdly, to harness the immense potential for learning about education that is to be found in its own formal administrative and evaluative machinery? Here the problem is that this machinery maps into wider authority and control systems that tend to preclude innovation, participation and the public practice of scepticism about prevailing explanations.

SOME METHODS BASED ON SEDA.

Our experience to date with the SEDA suggests ways of tackling each of these problems. It currently comprises some 350 variables on qualified school leavers in Scotland in the post-Robbins decade. We describe it in a separate, informal Conference session - Tuesday 7.30 to 8.30 p.m. By holding life-history data on pupils between, roughly, the ages of 14 and 19 years over a ten year period and on a local and national (Scottish) basis, we can begin to open particular practices and events to observation and evaluation in terms of wider parallels and consequences that lie beyond the ken of the single observer. This can, moreover, become a self-evaluation since the SEDA stores concepts and technical expertise in a form (the FILE system of the Statistical Package for the Social Sciences) that allows of their retrieval by the technically inexperienced. Since the archive can now be accessed through a near-national computer link, using a language that increasingly corresponds to standard English, there is no technical reason why anyone should not interrogate such concepts, refine them, supplant them, explore their correlates in theory-building or relate them to their local situation. Finally, the archive continues to enjoy mutually acceptable links with central government's statistical machine that allows one both to reproduce and to go beyond the official understanding of events. Similar links are currently being established with local government. This indicates that it is sometimes possible to harness the strengths of the formal machinery for monitoring and evaluation to produce a focussed and cumulative discourse that can, nevertheless, remain open to a variety of perspectives.

To develop this point briefly, the Scottish Education Department passes to the CES the names and addresses of the leavers who agree to co-operate with the CES research when once they have returned to the SED the information that it requires. The CES then collects additional data from these leavers. The SED also supplies information on population values which allows the archive data to be weighted to reduce response bias. Hence the many strengths of the official machine can be released for wider and more heterodox use without compromising what government takes to be its legitimate interests.

The concepts that are stored in the archive may then reflect either central or local understandings. 'Central' and 'local' refer not only to geographical areas but also to authority-, administrative- and belief-systems. Examples of centrally generated concepts currently include classifications of schools, curricula and of achievement used by such agencies as government, the UCCA or the Scottish Certificate of Education Examination Board. They enable one, broadly, to tell the story in terms of the persons who produce official versions of events and apply those categories and understanding to others in their practice. Locally generated concepts might include the application to the same raw data of further classifications reflecting, say, a headmaster's idea of what constitues a balanced curriculum or someone's beliefs about the type or quality of schools in his area.

Thus central resources and expertise can service local exploration and self-evaluation. These activities may, in their turn, make more explicit previously inaccessible practices and understandings. We may, for example, observe users' derivations of new variables from the raw data and the criteria by which they seek to validate them. Stored concepts can also be made to compete with each other in ways that make more explicit the deeper principles of educational practice: for example, what is meant by breadth in the secondary school curriculum, whether or not it is achieved and why. This, in turn, allows discussion of these principles or, where the grounds for a belief or practice have been insufficiently demonstrated, indicates more precisely where further information is required.

EXTENDING SEDA BY COLLABORATION

The Collaborative Research Programme will extend the SEDA's decentralisation of secondary analysis to the problem-definition and data-collection stages. The following are three of the methods we shall use in our attempts to achieve this:

(1) A survey of 1975-76 school leavers, comparable with our earlier 1970 and 1972 surveys, with a target sample of approximately 15,000 and an average 1 in 2 sampling fraction, varying from 1 in 5 for larger or less crucial categories to 1 in 1 for smaller or more crucial categories. We want large numbers, firstly, to permit analysis of local situations. Also, we wish to use perhaps four or five verions of the questionnaire with a core of information common to all versions. We shall control the content of the core. It will extend the SEDA time-series to 1976 and also permit comparability between the data collected in the several unique sections, or annexes, of the 1976 questionnaire in terms of leavers' life-history data. The content of the annexes to the common core will be decided by other persons who may include central and local government, academics of various theoretical persuasions and a variety of individual and institutional participants.

(2) The 1976 leavers' questionnaire will also be used as a base to which will be added later life-history data on subsets of the same leavers. Sources might include both institutional records and further data-sweeps. We shall not control the content of these data-sweeps except, perhaps within Edinburgh University. This design is particularly appropriate to socialisation and achievement studies. It is economical and more sensitive to local practice than a centrally administered follow up which tends to focus on the lowest, common, between-institutional denominator of behaviour and practice. More importantly, it extends yet further the range and variety of events in terms of which earlier educational practices and experiences can be evaluated.

(3) Data at the school, Census-tract, Divisional and Regional levels will be added to the 1976 individual-level questionnaire data. These will include administrative and environmental factors on which official data are routinely available. But we also hope that the opportunity of an evaluation in terms of effects on individual behaviour will encourage administrators and others also to test situations and practices that would otherwise remain hidden or untested. There is also obvious scope, especially at the school level, for the exploration of a variety of perspectives on school effects based both on academic theorising and on local knowledge. They may ent 1 further data-collection from the schools.

We see no reason in principle why we should not incorporate in the data-collection sub-designs that aim explicitly to improve or question the measures to which we and other participants might be committed.

In addition to its own substantive interests, the research team will service the wider use of the research and analysis facilities, providing all the necessary data-handling and documentation functions. It will also meet the financial costs of collaboration. A variety of approaches are currently being made to potential participants and it is hoped that a Standing Conference structure will emerge. This will aim to bring into contact with each other persons with a common involvement in problems, whether acknowledged or not. Areas of interest that such a Conference might elaborate through the research facility may well include the implications of devolution and regionalisation, the demand for post-compulsory education and examination reform.

SOME PROBLEMS OF COLLABORATIVE RESEARCH

But will it work? Here we reach questions that will occupy the evaluation component of the programme. This will concern itself, broadly, with the system's response to the technically enhanced possibility of knowledge. Why should people risk making themselves explicit, risk exposing their actions and beliefs to a wider, and possibly public, evaluation? Why should they respond to the idea of an 'educational system' with 'common problems'? Why should they concede a superior rationality to a social-scientific approach - indeed, to which of several? Justifiably or not, will people not prefer to ignore us?

Such questions in their turn raise a variety of issues to do with the role structure and epistemology of research. I will mention two here. They both concern the relationship between knowledge and control. Firstly, if people will not make available for study themselves and the situations they control, (whether classroom or county, individual or institution), how else shall we reach them? By deception? Through unobtrusive measures? By monitoring merely the face they present to the world? By the invocation of some academic right of entry and judgement? Indeed, can we reach them at all in any systematic and satisfactory way without their also actively embracing us, along with our methods and rationality? If we cannot, are we not deluded in theorising about situations that must be changed in order to be tested? And what differences then remain between the actor and the researcher, between the participant and the observer? Unfortunately there is not space to develop here the implications of this line of thinking for our current uncertainty over knowledge paradigms in educational research.

There is, secondly, a familiar and fundamental tension between knowledge and control that must be recognised. Briefly, theoretical advance requires us to erect structures for the long-term, extensive and systematic observation of educational systems. But to do this we require an authority - a capacity to command long-term compliance - that probably only government will ever, or should ever, enjoy. However, the maintenance of such authority tends to preclude a second pre-requisite of theoretical advance, namely the persistent, public attempt to falsify prevailing explanations. It is never easy for authority in any form to doubt its own account of events without compromising the executive effectiveness on which that authority rests. Thus we have a situation in which the major structures that might, by my argument, promote public learning, cannot easily be cast in this role. Nor can they ever be as long as we give authority only to those who claim to know, and not to those who doubt.

The Collaborative Research model is, then, more than a tactical ruse to squeeze a research quart out of a conventional pint bottle or to engage official actions and understandings in critical discourse. It is more than an attempt to make results more relevant or dissemination more effective. It is, in addition, an epistemological position that acknowledges the central theoretical relevance of issues of resources and power that academics commonly dismiss as a tiresome and theoretically irrelevant matter of means. It holds that any unilateral attempt by 'we researchers' to understand 'those educationists' will always fall short of our theoretical ambitions because of the depletion of resources for understanding that results from such distinctions. To relapse for a moment into exhortation, it means that we must include an account of our own limitations - a theory of research practice - in our understanding of the system; and also that we must recognise the theorist in all of its actors. Its most immediate practical implication is that servicing and teaching are central to the research activity - that we shall best explain the world by enhancing its capacity to explain itself.

But, to repeat the question, will it work? Will people collaborate? Obviously I must temper scepticism with enthusiasm. I would like to think that relatively powerless people will want to ask questions when there is the possibility of an answer; that authority would want to be let off the hook of always being expected to know; and that claims to authority based on the assertion of knowledge cannot help but be open to exploration. It depends, I suppose, on whether there is any force in an argument. If not, then at least it is all data for the evaluation. No programme should be without some form of collaborative research.

[1] The SSRC is funding the Scottish Education Data Archive project (HR 3187) and the Collaborative Research Programme (HR 3262). I am also grateful to the Scottish Education Department for continuing financial support and co-operation. The SED does not necessarily share the views expressed in this paper.

SOCIAL AND HISTORICAL CONTEXTS OF EDUCATIONAL RESEARCH

GODFREY THOMSON AND THE DEVELOPMENT OF PSYCHOMETRICS IN SCOTLAND, 1925–1950.

H. M. Paterson.
University of Glasgow.

My thesis in this paper is as follows: psychometric approaches to educational problems flourish in periods when the stress is on increasing the efficiency of the educational system in bureaucratic and rationalistic terms. In other words, a quantified and differential psychology which rests on assumptions about the measurement of psychological variables (in this case, the measurement of "ability" or "intelligence") will always grow in a climate where the most rational use of scarce educational resources is the paramount concern. Psychometrics becomes, first, the tool of a bureaucratic bloc (the administration) intent on rationalising a maturing educational system; and, second, a major component of the rationale for the bloc's organisation of the system itself.

What I want to do in this paper is to try and explain the rise of psychometrics in Scotland, between 1925 and 1950 in terms of these ideas. There was undoubtedly much psychometric activity in the country before 1925 but my focus is on what I take to be the leading school of educational psychometrics which was to be found in the Department of Education at Edinburgh University and in Moray House Training College. Both were dominated by the figure of Godfrey Hilton Thomson, Professor of Education in the University and Director of the College, a joint appointment which he held from 1925 until his retirement in 1951. While Professor of Education at Armstrong College, Newcastle, Thomson had been invited by Northumberland Education Authority in 1921, to investigate the reasons for a relative scarcity of rural children in the county gaining "scholarship" places in secondary schools. In response, Thomson developed the Northumberland Tests which were the first group tests of ability to be used on a large scale in Great Britain. In constructing them, he was plainly influenced by the U.S. Army tests which had become internationally known after 1918. From then on, a major focus of his work became the identification of children of high natural ability who might have been missed by the usual tests of school attainment but who were nonetheless intelligent enough to benefit from an academic course of secondary schooling. Even Thomson's most hightly mathematical work in factor analysis derives from this source - of factor analysis, he writes:

> 'It is probable that the subject-matter of this book may seem ... to be far removed from contact with the actual work of schools. I would like therefore to explain that the incentive to the study of factor analysis comes in my case very largely from the practical desire to improve the selection of children for higher education. When I was thirteen years of age and finishing an elementary school education, I won a 'scholarship' to a secondary school in the neighbouring town ... I have ever since then been greatly impressed by the influence that event has had on my life, and have spent a great deal of time in endeavouring to improve the methods of selecting pupils at that stage and in lessening the part played by chance..."

The academic and "scientific" study of education and psychology was already well advanced in Edinburgh before Thomson arrived there. But Thomson wanted to take this much further by developing a leading school of educational research which would deal exclusively with the rigorous application of statistical and mathematical techniques to educational problems.

Of psychology he knew little at first, but he proceeded to take the University of London Diploma in Education externally, and in 1911 he enrolled in Charles Myers' psychological laboratory in Cambridge to study psycho-physics under the influence of William Brown's work. Brown had published The Essentials of Mental Measurement in the same year, devoting the first part of the book to psycho-physics but expounding the work of Pearson and Spearman in the remainder. It was this correlational work that was to really excite Thomson, on which he came to concentrate and which he was to advance, and he was soon to collaborate with Brown in producing revised editions of the latter's book.

It is clear that Thomson brought with him, from the natural sciences, a set of views as to the nature of scientific method and a set of techniques which he believed were more or less directly applicable to the problems of the human sciences. He was aware of the differences to some extent, but nonetheless the natural sciences represented an ideal which the human sciences ought to emulate. As he puts it at one point;

> "much mathematical study and many calculations have to precede every improvement in engineering, and it will not be otherwise in the future with the social as well as the physical sciences."

or again, referring to his own Sampling Theory of mind:

> "this theory is preferred because it makes fewer and less special assumptions, because it is more elastic and wider, and because it is in closer accord with theories used in biology and in the study of heredity. (my emphasis)"

Thomson believed that educational research was most likely to be made scientifically respectable through the application of such techniques in the shape of educational psychometrics, to the identification and quantification of psychological variables important to education. The single most important variable, in Thomson's eyes, was "intelligence", and he devoted his life-work to devising, first, increasingly sophisticated instruments for its measurement and, second, a mathematically coherent theory of functioning intelligence.

Thomson seldom spoke or wrote on social or educational matters outside his University classes, and so we are restricted for evidence about his views on them to those few of his writings which deal with such issues directly, and to scattered comments elsewhere.

In summary, Thomson' good society is a stratified one since the division of labour is a necessary prerequisite of social efficiency, and social efficiency is what enables the group to survive in competition with other groups. This involves a fine differentiation of society's talent, and thus purposive and rational educational selection replaces the harsh and chancy business of natural selection. The education system, then, becomes the major social instrument for the identification and fostering of ability; to allow bright children to languish in schools or courses for which they were not suited was to waste a scarce national resource; conversely, to allow "mediocre" children to follow academic courses was to do them, and the country, a disservice. More rational and more organised allocation of talents to courses was therefore a national priority. Such views require that talent be identified - and identified in such a way as to avoid possible charges of bias on the part of those who do the identifying. The only way to achieve this with sufficient accuracy and at not too great a cost was to quantify the notion of "ability" and to construct means of measuring it in individuals who could be treated in the mass. The U.S. Army tests, the Northumberland tests and the Moray House Intelligence Tests were all designed to do exactly this. The new "science" of psychometrics was to develop a technology of testing which could be used to calibrate precisely and "objectively" the needs of children, the needs of the educational system, and the needs of society, and would thereby ensure that these connecting parts meshed more smoothly.

By 1925 - the year of Thomson's appointment to Edinburgh - the rational development, along bureaucratic lines, of the

Scottish educational system was well advanced. Scotland had had what amounted to a nation-wide system of schools for quite some time, and it is probably not an exaggeration to claim that the Education (Scotland) Act of 1872 mainly confirmed the legal existence of a de facto framework which covered most of the country and which was supported financially and conceptually by its people. There was also a well-developed structure of local and central government ready to take over the administration of a national system of schools and this helped to ensure that, by 1872, the legislation which constituted a radical and difficult innovation in England was in Scotland relatively easy to put into practice.

These were the beginnings of the bureaucratisation of Scottish education and, in the years to follow, this feature was to become marked. Central control of Scottish schools by the Education Department was to be the norm from the very beginnings of the national system; by the 188O's the Department was "a powerful bureaucracy set in its ways" (Lenman and Stocks, 1972), dominated by the figure of the Permanent Secretary. These were no rubber-stamping civil servants, channelling the edicts of Whitehall to the schools of Scotland, but men with clear ideas of how Scottish education should be run most efficiently and with the determination and power to see these ideas translated into official policy and practice. The influence of such bureaucratic administrators as Craik and Struthers is partly to be explained by the fact that Scotland has never had a Minister of Education since education is simply one part of the multiple brief of the Scottish Secretary of State, who tends therefore to delegate much policy-making to top civil servants or Junior Ministers. Between 1872 and 1914, the efforts of the Department (via the local School Boards) were mainly directed to the development and rational organisation of the elementary schools. But after the war, the Scottish Act of 1918 brought an important shift of emphasis. The 9OO-odd local School Boards disappeared, to be replaced by educational units of greater size and efficiency dealing with whole counties or burghs and not the smaller parishes. These were "advised" (and advice from the Department was not lightly ignored) to appoint an executive officer - soon to be widely known as the Director of Education - who could give expert advice to the new authorities. Thus, a further step was taken in the efficient rationalisation of the system at local level; an expert was placed at the helm of local governance of education, to correspond with the similar placing of expertise in central governance; and expert might now talk directly to expert without too much intervention by concerned (but amateur) outsiders.

The Act was followed by a Departmental Circular No. 44 in 1921 which dealt with the re-organisation of secondary schools and the transfer of children to them from the elementary schools. This Circular made it clear that the view of the Department was that there should be two separate forms of secondary school - one for the talented and able who would be likely to carry on to university, and another for the great bulk of the population who would not do so. The intention was laudable; the secondary education of the great majority of children had been badly distorted by the bookish and academic tradition of such schooling in Scotland. But the consequences were dire. The Circular was pushed through by the Scottish Education Department against much opposition, and firmly established the principle of different schools for different abilities as official policy. Such a step meant that the efficient grading of children into two groups of ability became a prime administrative problem, and thus the identification, selection and allocation of ability at the age of "transfer" was officially established.

This was no new departure for Scottish education, nor was it foreign to the Scottish educational temper. An essential ingredient of the Scottish educational tradition was the way the school system was available to the whole population, from whatever background or social class. In this sense, there always has been a strongly egalitarian tinge to Scottish educational policy-making. At the same time, however, this egalitarianism is compounded with an emphasis on the "natural selection" of ability whereby many children are given a first formal chance of academic schooling but only the "fittest" survive the course. Thus the "lad o' pairts", who might come from any section of Scottish society, was the child whose intellectual ability was good enough, and whose moral qualities were strong enough, to see him through the academic trials awaiting him on his rise from parish school to the university and the professions. Those without the necessary "pairts" were left to fall by the wayside - and Scottish secondary education today is still characterised by a very heavy wastage of children who depart at the statutory leaving age. In this way, Scotland has managed to keep the schools open and equal by admitting a relatively large percentage to academic secondary education, but also more efficient and closed as the educational ladder is ascended. In other words, the Scottish system manages to combine egalitarian and selective emphases in a tradition where the use of the system's ladder to foster individual upward mobility in a competitive situation is, as McPherson (1973) points out, quite taken for granted.

Circular 44 of 1921 is also important in that it shifted the responsibility for constructing schemes of selection from the Scottish Education Department to the new local authorities and their Directors of Education, who were to submit such schemes for the Department's approval (thus getting rid of the chore while keeping the power of veto). Thus the Directors were faced with the administrative and technical problem of ensuring that the increasing numbers of children passing on to the secondary schools were properly sorted and allocated, and of doing this as efficiently and as cheaply as possible in times of some economic malaise. The identification of the able had officially begun almost thirty years earlier; in 1892, the Education Department had introduced the "Merit Certificate", a test of attainment for thirteen-year-olds at the satisfactory completion of their elementary schooling which effectively identified those who passed as being "fit". Within six or seven years however the Merit Certificate had ceased to be merely such an identificatory device and had become instead a passport to the equivalent of secondary courses at the time. It had become a means of selection which stated that those who passed it were fit to go on to advanced work. But by the 192O's, the problem of allocation was becoming crucial since children were no longer to be simply promoted to secondary schools but were to be directed to those courses which were best suited to their level of ability. In attempting to solve this problem of identifying, selecting and allocating children according to their ability, most Directors drew up schemes which relied initially on tests of attainment (mainly in English and Arithmetic) and the teachers' estimates of likely future academic success (or failure). But there was a clear recognition of the disadvantages inherent in these measures. Tests of attainment might merely reflect amount of time devoted to the tested subjects in the elementary school, which, in turn, could (and did) lead to severe distortion of its curriculum. Teachers' estimates were bound to be subjective and open to all kinds of possible bias. Neither attainment tests nor restimates was a sufficiently "rational" or "objective" way of assessing potential, and, as parents increasingly realised the effects of selection on the life-chances of their children, the possible bias of such measures became increasingly difficult to explain away. What seemed to be required was a means of assessing the innate ability of each child which would be as respectably "scientific" (and thus objective) as possible, but which could also be easily and cheaply administered to large numbers. Such measuring instruments would meet the demands for justice in selection, for efficient and rational selection, and for cost-effective selection. The Moray House Tests of Intelligence, which Godfrey Thomson and his colleagues and students were to develop in the twenties and thirties, were to offer all this to hard-pressed administrators in Scotland and elsewhere.

Thomson's attitudes, then, to the functions of an education system correspond closely to the official policies of the administrative bloc in Scottish education. They all shared a belief in selection as a necessary means of producing an educational elite; and they agreed that the need was for "scientific" instruments which would differentiate the abilities of children so as to properly guide these children into secondary courses suited to those abilities. Small wonder then that the administrative bloc seized on group tests of ability as an answer to their administrative difficulties. Such tests were designed to quantify ability (and thus identify it in obvious fashion) as cheaply and as efficiently as possible, and they therefore greatly eased the purely technical problem of allocating large numbers of children to a variety of secondary courses. By 1933, almost forty percent of Scottish local authorities were using intelligence tests as part at least of their "promotion (or transfer) procedures" for allocating children and the practice was to become universal by the 194O's when the impact of

McLelland's thorough and detailed work on Selection for Secondary Education (1942) was great.

The institutional validity of psychometrics in Scotland was further guaranteed by Thomson's work in developing a flourishing programme of teaching and research at Edinburgh. His attempts to make educational research more rigorous by the application of mathematics and statistics found a receptive audience in Scotland between the wars, so that it was not long before the Edinburgh Department and Moray House became the place in which to study educational psychology and in which to learn the craft of conducting educational research "properly". He succeeded in attracting able and well-motivated post-graduate students, many of whom proceeded after their degree to high posts in the educational and administrative hierarchy. The quality of Thomson's teaching was justly famous, to judge by reports, but mainly in his explanations of the intricacies of test construction and factor analysis to a relatively non-mathematically-minded group of students. The coherence of the theory of tests, and analaysis of the consequences of using tests in school selection, seem to have been left largely unexplored - although some of his students were always aware of the narrowness of his particular view of educational research. Although such a view was never totally accepted, Thomson came to dominate much of the educational research work being done in Scotland - he was, for example, a founding member of the Scottish Council for Research in Education and his influence here was considerable, although he himself seems to have been directly involved only in the two massive national Mental Surveys of 1932 and 1947.

Such dominance was not to last. The growing criticisms of early selection and of the indiscriminate use of group mental tests reached a crescendo in the 1950's. The ever-increasing amount of sociological evidence that the "objectivity" of testing was fallacious helped to strengthen the attitude that testing was more a bureaucratic convenience than a "scientific" process for determining what sort of education one's children were to receive. In addition, the growth of egalitarian social attitudes after World War II led directly to the political adoption of ideas on comprehensive schooling which eschewed the early segregation of children. The public legitimisation which psychometrics had enjoyed since the mid-1920's was largely withdrawn.

Thomson's career is perhaps a good example of the dangers and delights of close links between educational theory and research, and educational policy-making. When the administrative bloc of Scottish education were intent on developing further a national system of schooling, they required refined instruments for the identification, selection and allocation of children to fit that system. When the stress was thus on efficiency and coherence in developing a rational system of schools the quantified and seemingly objective results of psychometric procedures provided a respectably "scientific" justification for the system, as well as providing a ready-made set of tools to be used in setting it up and in furthering its growth. But the standing and influence of psychometrics was to wane with changes in public attitudes to its educational consequences, and political responses to these changes.

For their comments on an early draft of this paper, and for information and inspiration, my thanks to: Bob Bell, Dave Hamilton, Richard Hamilton, Malcolm Mackenzie, Stanley Nisbet, Douglas Weir, Eric Wilkinson and all participants in a Department of Education seminar in Glasgow University.

(For a complete draft of this paper, with notes and references, please apply to the author).

GODFREY THOMSON AND SCOTTISH EDUCATION

(SUMMARY)

R. E. Bell.
The Open University.

The first of these papers has rightly celebrated Thomson as a major figure in the development of British psychometrics and as one of the founding fathers of British secondary school selection. Clearly, the major achievement of his working life was the perfecting of techniques for such selection and, in particular, the development of the Moray House Tests as major items in the post-war selection procedures not only of most Scottish local authorities but of the majority of English and Welsh ones also. At the same time, Hamish Paterson has rightly warned us against assuming that Thomson could therefore willingly embrace the whole Black Paper deal. Robertson - and Vernon more specifically - reminds us that Thomson's ideal secondary school was more akin to Dewey's Community High School than it was to the post-war, elitist, English grammar school (from which the old leavening of less intelligent fee-payers had been removed, amid grammar school opposition, after 1944). He believed in efficient testing as a means of allocating students to suitable courses but preferably not to separate institutions and he thus tended to favour the Scottish multilateral school rather than the more socially divisive English system. Moreover, Thomson foresaw, as Robertson again reminds us, that sociology might soon overtake psychology as the most pertinent discipline in the formation of secondary school policy, and he might well have been less disturbed than Cox, Dyson and Boyson by the present influence of Halsey, Bernstein and their colleagues. For him, to describe a child as "intelligent" was to describe it and no more. It certainly involved no assumption of a right to social superiority or a right to superior educational treatment. Nor did he value highly the English emphasis on specialisation either in teaching or learning.

It is significant that despite his later reputation as a teacher of statistics and psychometrics and as the "founder" of Edinburgh's reputation in those fields, he neither introduced such subjects as major items in the Edinburgh (or Scottish) curriculum nor did he choose to give lectures in those subjects during his first years as professor. He apparently preferred to build up the body of his philosophy lectures and it was his "Modern Philosophy of Education" that proved to be his major publication during the first five years of his period at Moray House. Thus Dewey appears to have loomed far larger in his early teaching programme than did the new American and London psychologists and it was appropriate that Thomson should be present on the platform on the 4th May 1929 when the Educational Institute of Scotland conferred upon Dewey the degree (sic) of Fellow. Also present was Boyd, the non-professorial head of the Glasgow University Education Department and this was one of the few occasions when these two celebrated figures, who concurrently headed the major Scottish university departments for some twenty years appear to have co-operated in public - significantly on a purely ceremonial occasion. Stereotyping has made Boyd the gentle humanist and Thomson the stern statistician, though their preoccupations as researchers and teachers suggest that, as over Dewey, they had far more in common than is usually supposed.

William Boyd is, of course, still a relatively well known figure outside Scotland. His "History of Western Education" has recently been considered worthy of reissue in a revised version by Edmund King while his handbooks on Plato and Rousseau are still standard textbooks in universities and colleges; and this fame in the historical/philosophical field has tended to divert attention from the fact that Boyd and his colleagues in Glasgow always saw themselves as just as devoted as Thomson to the cause of experimental psychological research and what was usually called in Scotland, even under Thomson, and his successors in Edinburgh, "Experimental Education". Moreover, in

Glasgow Boyd and, later, R.R. Rusk, developed and used psychometric tests which may not have stood the test of time like those of Moray House but were at one time just as widespread and influential among Scottish teachers. Indeed, Boyd and Rusk openly mistrusted some of Thomson's approaches which they saw as too impersonal and pedagogically restrictive and for many decades, Rusk (who died in his nineties) continued to attack Moray House Tests, with, according to many eye-witnesses, "tears in his eyes". Nor can Rusk be dismissed as an old-fashioned insignificant Scottish figure. He shared Thomson's academic background of the new London school and of Germany and, in addition, had studied psychology as a post-(Glasgow) -graduate undergraduate in Cambridge (where, incidentally, he claimed, through his work as a correspondence course tutor, to have concurrently given Bartlett his initial grounding in psychology!) Moreover, both he and Boyd were much more closely identified with the mainstream of Scottish teachers than was the rather more Olympian Englishman Thomson. He, like his two successors (also both Englishmen), tended to hold aloof from Scottish educatinnal politics and particularly from the Educational Institute itself.

The importance of the Educational Institute of Scotland in the development of Scottish educational research has usually been underestimated. Its recent preoccupations with salaries and conditions of service have led it to emphasize its trades union role at the expense of an earlier role which once distinguished it more clearly from the English National Union of Teachers. For, initially, its self-image had been that of a learned society and, unlike the N.U.T., its membership had embraced the full range of Scottish teachers, including not merely elementary teachers (the only ones to be unionized in England) but "schoolmasters" in the new Victorian private sector and the University professors themselves. Its charter allowed it to confer degrees most of which were honorary but some of which were earned by thesis. As a result, even when it was actively exercising its trades union functions, its notions of professional status caused it also to nurture continually the expansion of its academic activities.

The years following the first world war were crucial years in the professionalization of British teachers and, in Scotland, the need to encourage the higher study of education and of research among teachers themselves often dominated the E.I.S's rhetoric if not its financial priorities. Indeed, the perennial cry of Scotland betrayed was often linked to such issues. Many Scottish teachers were aware that wartime experience had given a massive fillip to psychological and thus to pedagogical and psychometrically orientated educational studies in England and America.

Up to 1918, membership of the British (and at that time almost exclusively English) Psychological Society had been less than a hundred (98 in the Annual Report for 1918). Its meetings were leisurely, the minutes usually recorded the restaurant to which they adjourned, and the papers were often physiological or philosophical in an earlier tradition. Papers dealing with specifically educational interests had only begun to creep in after 1910. Valentine, then at St. Andrews, had been elected in 1912, Rusk in 1913 and Thomson a year later. But in the immediate post-war years, the situation changed. Demobilized R.A.M.C. officers and crash-trained occupational psychologists, who had run the Army selection procedures, increased the membership by many hundred per-cent (631 by 1920, of whom 366 were described as "educational") and grants began to be made by the society for specifically educational research. In 1923 a tiny Scottish branch of some 8 members was formed under Drever (which, interestingly, included neither Boyd nor Rusk) and it was mainly based in Edinburgh. Its meagre membership reflected how far behind Scotland had fallen.

Although all four universities had instituted a post-graduate degree (Bachelor of Education) in education and psychology by Ordinances passed during the war, by 1920 the Ordinance had been activated only in Edinburgh where a handful of candidates (including some from overseas) had begun to study for it.

Against such a background, it is not surprising to hear the President of the E.I.S. in 1919 drawing attention to the new posts of Director of Education (already discussed by Paterson) and the lack of Scottish attention to their training.

In order to ameliorate this situation, the E.I.S. began (unsuccessfully) to press once more upon Glasgow and Aberdeen the need to establish chairs of education and Boyd, as President-elect, suggested research should be the major topic at their 1920 annual conference. In preparation, on the 1st. November 1919, an Educational Research Committee of the Institute met for the first time. In his recent history of the S.C.R.E., Craigie suggests that any meaningful account of the transactions of this Committee (a precursor of the S.C.R.E.) have been lost. In fact, however, the general minutes of the Institute contain quite detailed records of its proceedings. Its first membership included not merely Boyd and the Glasgow group but McLelland (then a county Director) and Drever and his Edinburgh colleagues, who had already established the basis for that experimental work often assumed to have been first established by Thompson. Nevertheless, Boyd remained the most enthusiastic and persistent member of a committee whose zest and activities tended to fluctuat e. His scientific skill may now be questioned but his devotion to the cause of testing and "experimental education" was considerable.

This is of significance, for when the Glasgow Ed.B. Ordinance was eventually activated his organisation of study there, on a part-time as well as a full-time basis, meant that his eventual recruitment of students far out-stripped that of Edinburgh. When the present writer carried out a questionnaire survey of all traceable holders of the Scottish education degree in 1968-69, Glasgow graduates (according to University records) numbered 409 compared with Edinburgh's 281.

A myth has arisen that Boyd's major preoccupations in teaching his Ed.B. students were philosophical, in contrast with the "scientific" preoccupations of Thomson. In fact, however, of the theses presented for the degree in the period before 1950 when the influences of Thomson and Boyd were at their height although some 83% of the Edinburgh work was on experimental psychology, testing or similar topics, 60% of the Glasgow theses were also on such topics while a further 26% took the form of statistically-based curricular surveys of various kinds, rather than philosophical/historical theses which were extremely rare in either university before the late 1950's.

It is of interest to notice also that there was no great difference in the eventual careers of the Edinburgh and Glasgow graduates. The longest jobs ("longest" as opposed to "first" or "last" job because the former may well have been a temporary expedient and the latter a retirement job) undertaken in their post-B.Ed./Ed.B. period for the pre-1950 graduates appear to have been as follows:

TABLE 1

	Edinburgh	Glasgow
1. Primary/Secondary teaching	18.9%	26%
2. University/college lecturing	32.1%	32.7%
3. Educational psychology service	22.6%	24%
4. L.E.A. officials, inspectorate	22.6%	11.5%

The differential in the case of the fourth category may well reflect the eventual attraction to an employing authority of an intimate knowledge of the Moray House testing techniques. On the other hand, in terms of influence on Scottish education, it is worth bearing in mind that Glasgow graduates were more likely to work in Scotland.

In any case, the nature of the contribution of the Boyd degree (supervised between 1945 and 1950 by Rusk) is not so dissimilar to that of the Edinburgh degree as is sometimes supposed. Boyd was a pioneer of child guidance clinics. He opened the first such clinic in the West of Scotland and actively involved his students in it. This meant that his preoccupations in the testing field tended to be as much concerned with the diagnosis of individual problems as with selection, though it would be wrong to suggest that both he and Rusk were not also preoccupied with the efficient performance of the latter. Moreover, he shared this interest in clinical work with Drever who ran his own clinic in Edinburgh and whose influence over individual B.Ed. students was often as strong as that of Thomson himself. Our survey evidence suggests that for much of the time Thomson and Drever appeared (to students at any rate) to communicate little. Although he had been a Scottish pioneer in psychometry, Drever's main

interests began to lie more and more in the fields of occupational psychology and clinical work, though he shared Thomson's aversion to over-specialization and had begun his academic career as a lecturer in comparative education. Thus he and many of the apparently Thomson-orientated graduates may well have shared the enthusiasms of Boyd and Rusk rather than those of the Moray House psychometrists, whose claim to form the central core of Scottish research work in the pre-1950 period must be treated with considerable scepticism.

Certainly Drever cooperated with Boyd in the E.I.S. research committee which, at the start, thought less in terms of mass surveys and nationally standardized tests than about the encouragement of research by individual teachers and the circulation of information on experimental psychology. However, testing did soon begin to occupy their attention. It was important not merely to the new Directors of Education (to help whom the committee had been formed) but was of obvious importance also to the clinics and diagnostic tests in arithmetic and composition were projected at the first two meetings. Boyd's trial of these in a sample of self-selected schools hardly inspired confidence in Thomson's future disciples and, significantly, attendance by Edinburgh and Aberdeen members began to fall off. More diffuse projects such as one for spelling reform were mooted, and there was talk of "research" prizes for essays by teachers. Nevertheless, as early as March 9th. 1921, proposals were submitted for a "simple test of general intelligence by means of vocabulary for use at the qualifying stage" and during the next few years group testing was often on the agenda. At a meeting in 1923, for example, the committee hoped "to take up the question of the educational value of Intelligence Tests" and, in June 1924, the E.I.S. executive was asked to institute a general survey of the means being used for "promotion of pupils at the qualifying stage". By the 27th. March 1925 Boyd had drawn up a battery of tests which the E.I.S. was to suggest for this purpose, and in December of the same year, when Thomson had hardly begun to play himself in at Edinburgh, plans were also made for an Intelligence Test to be used "at the Leaving Certificate stage".

Over 30 (self-selected) schools agreed to take part in this experiment and on 4th February 1926 there is the first mention in the minutes of a larger-scale "National Organisation of Research in Education", discussed in a memo by Steel, headmaster of Allan Glen's School in Glasgow. This discussion pre-dates by a year the first meeting with the Association of Directors of Education which finally led to the establishment of the Scottish Council for Research in Education, which was from the first a tenant and protege of the E.I.S. Significantly, at the eventual meeting in December 1927 from whence the Council sprang, the Chairman acknowledged the Research Committee's primacy as the originator of the idea. Moreover he linked the move to establish the Council more to the Directors' need for guidance on pedagogical and curricular issues of the sort that interested Rusk and Boyd than the need for guidance on selection procedures. Nevertheless, given the growing importance of selection and the fact that the new Council's primary aim was seen to be the servicing of Directors' needs, it was not surprising that questions of selection should early occupy their attention in the short term.

Significantly, Thomson's contribution to this inaugural meeting was concerned with the need to inaugurate mass surveys, which only a national organisation could handle, and the need to protect the personal and professional rights of research workers vis a vis the Council, rather than with the specific problems of Directors and the daily needs of the classroom.

By this time the EIS research committee had blossomed into a body which, says Vernon, "put English educational researchers to shame" and to which, as the SCRE report for 1936-37 rightly boasts, there was no equivalent body elsewhere in the British Isles.

There is no doubt that the uniqueness of this creation must partly be related to the preoccupations of the EIS itself, as a body embracing all sections of Scottish schooling. Significantly, the prime mover from their side in the establishment of SCRE was the headmaster of a relatively elitist fee-paying school - a man who would never have become a leader of general teacher opinion in the more divided and class-ridden English profession. Moreover, there is similar significance in the fact that the university departments primarily responsible for the training of research workers in Scotland - through the education degrees - were, from the first war onwards, not concerned, like their English counterparts, with the certification of specialist grammar school teachers - for such people were trained by the Scottish colleges, so that the universities were left free to think about educational studies in more general terms.

On the other hand, despite the Universities' warm welcome to and involvement in SCRE, its new part-time director was not a university figure congenial to Thomson but a college figure, Rusk, who was mistrustful of Moray House and all its works and whose books were no longer used for teaching purposes in the Edinburgh department. As we noted earlier, Rusk had been the real pioneer of "experimental education" in Scotland had had published an introduction to the subject as early as 1912. Significantly, however, while this and later volumes covered the topics which most concerned Thomson, they also covered straight pedagogical issues such as reading and aesthetics, to which the Glasgow department (and the Edinburgh psychology department) paid far more attention. Astonishingly, in view of their cooperation in SCRE over many years, Rusk's "Outline of Experimental Education" (1960) does not even mention Thomson (or Moray House Tests) at any point. The fact is that despite his national influence in relation to secondary selection and his distinction as a statistician, Thomson played a far less dominant role in the SCRE and Scottish research generally than those who always see him as a giant among pygmies might suppose. Certainly his work on testing was infinitely superior to that of Boyd, but Boyd's tests were influential and widely used among teachers, while much of SCRE's remit did not interest Thomson at all. As he himself said, he had "a flair for inspiring and conducting big surveys" and this influenced important SCRE work but he had little interest in their major, pioneering investigations of the curriculum and in their proposals for widely based cohort studies such as that of university entrants proposed at an Eastbourne conference in 1932 which anticipated the sociological-cum-attainment surveys of forty years later.

There was then a sense in which Thomson's work, far from being central to the main body of Scottish educational research, stood a little aside from it. A geographical analysis of his B.Ed. graduates is revealing. Whereas 97.2% of pre-1950 Glasgow Ed.B.'s had graduated first in Scottish universities (and only 1.8% outside the British Isles) only 66.4% of Thomson's graduates had earlier graduated in Scotland while 13.1% came from outside these islands. Similarly, while 77.9% of the years worked by pre-1950 Glasgow Ed.B's after their graduation have been worked in Scotland, only 52.3% of the Edinburgh work years have been spent there. 30% have been spent in other parts of the U.K. and 17.7% outside. (The comparable Glasgow figures are 15.1% and 7%). Thus Thomson's preoccupation, like that of Edinburgh university in general tended to be as much with the world outside Scotland as with internal problems.

Moreover, despite his undoubted and accepted distinction as a scholar at the time of his retirement, it would be wrong to assume too readily that Thomson was initially welcomed by Edinburgh University as a distinguished scholar and a potential adornment of the university in the first place. For the truth is that he performed against a general Scottish background of intolerance of educational studies within the university sector.

It is true that Edinburgh and St. Andrews had been the first British Universities to have Chairs of Education, but with the growing elitism of even the more far flung universities after 1900, educational studies, whose social academic status was generally low, often came off badly. It is difficult to see the combining of the Edinburgh chair with the headship of Moray House Training Centre, (a temporary arrangement only operative during Thomson's tenure of the post) as anything but a device by certain Edinburgh professors to discredit educational studies (along with its professor) and to save expense. Moreover there was an undoubted political movement within the university which always opposed Thomson, and which succeeded in denying his post and the Moray House empire to a suitable psychometrist on his retirement; indeed, his work met with emotional disapproval in a wide academic circle.

Robertson has suggested that "Scotland's recognition of Godfrey Thomson (was) slower and less generous than it should have been.". Certainly he created enormous enthusiasm in certain students and great respect in many

others; certainly, as an individual, he brought to educational studies in Scotland a better scientific example than they had generally known earlier. But possibly Robertson under-estimates the extent to which Thomson stood aside not just from Scottish life in general but from her teacher life, as represented by Boyd, Rusk and the EIS in particular. Like Paterson, Robertson probably also exaggerates the centrality of his thinking, not just within the Scottish educational studies of his time, but even among those who studied for the Edinburgh B.Ed. It may be that he is the most important single scientist who has been concerned in Scottish educational life, and that the perfecting of Moray House test standardization, to which Vernon pays such a warm tribute, may be the single most rounded and satisfying achievement of Scottish educational research so far - but neither Thomson nor the Moray House tests may prove to be as lastingly central as his disciples usually suggest. As Vernon again suggests, he may have had very little long-term influence on Scottish education as such, not because secondary selection has now been swept away but, ironically, because he was too liberal to take the preoccupations of Scotland's academically-minded teachers seriously. The scientifically less satisfactory world of Boyd, Drever and the EIS research committee responded more readily to specific teacher needs and may continue to influence us even now, not just through SCRE and the clinics but also through Boyd's students. Thomson's standards and techniques may still, rightly, influence many Scottish educational scientists but the direction and subject-matter of his research have now to a large extent become irrelevant and lost.

(Summary: for a complete draft, plus notes and references, please apply to the author).

KARL PEARSON AND STATISTICAL EMPIRICISM

B. Norton.
University College, London.

This brief and informal talk has two main parts.

The first and larger part seeks synoptically to explain how it was that Karl Pearson (1857-1936) came to make those contributions to statistical theory which Churchill Eisenhart has described as having 'established statistics as a discipline in its own right', and which Victor Hilts has suggested were eventually to change the character of almost all social science'. After indicating the roles of Adolphe Quetelet (1796-1874) and Francis Galton (1822-1911) in the history of statistics, the talk suggests that Pearson's own development of statistical theory was a necessary condition for the progress of his and the zoologist Weldon's new, would-be discipline of biometry: a discipline best described as an approach to mathematical evolutionary theory, in which the accepted problems and acceptable modes of solution for them were laid down in a very distinctive manner. (Note, for example, that when Charles Darwin spoke of natural selection he described it as 'this preservation of favourable individual differences and variations, and the destruction of those which are injurious', whereas, for Pearson, it was that which 'is measured by the changes due solely to mortality in the mean and standard deviation of the variation curve as we pass from one adult generation to another'.) It is then further suggested that Pearson's desire to develop this new discipline in its characteristic form arose from his deep commitments to, on the one hand, a radical Darwinian view of man, society and morality, and, on the other, to a set of distinctive views about the nature, purposes and aims of science - both acquired during his early life, particularly during his spell as a philosophy student in Heidelberg in the late 1870's. At this point, it will be argued that Pearson saw the new statistics as a new mathematical methodology for the sciences, notably for the social sciences, which, in a strong sense, encapsulated the views about science mentioned above.

The second and smaller part discusses Pearson's actual practice in the biosocial sciences and indicates points arising from it, its relevance, it is hoped, is fairly general and far from unique to Pearson's work. Two points will be dealt with, (i) questions which concern the aims and goals for science which we have to postulate in order to make sense of actual historical practice, and (ii) issues concerning the relation between 'integrity' and 'worthwhileness' in the social sciences on the one hand, and discussions of what sort of a society we ultimately want on the other. (Summary)

"TWO PARADIGMS OF EDUCATIONAL RESEARCH?"

George Hudson.

While educational research in England and Wales experienced a late start compared with other industrialised countries as an organised activity, today it is well established as a distinct activity. In the U.K. there are three national organisations and a number of university departments devoted to empirical research in all sections of our educational system. A notable number of people, myself included, are now using a multi-disciplinary approach to examine, analyse and make pronouncements on the state and workings of education.

My reason for writing this paper is that we earn our living by using the wealth created by others and the time and energy of people - the teachers and pupils. Our product as researchers is knowledge, which is abstract and certainly cannot be exchanged directly for the weekly groceries. We hope, therefore, that the knowledge of educational research will contribute to the general good of society and education in particular. Failure to do this means that research is an exploitive rather than useful activity. The traditional approach for justifying and legitimising educational research is to catalogue a list of research findings and hold them up to the public as useful, insightful contributions to the general good. The validity of knowledge cannot be established by blowing this trumpet of self-justification but by examining the grounds upon which it is based. The grounds for knowledge is based upon the assumptions, beliefs and values of researchers, what Kuhn defines as a paradigm:

> "... the entire matrix of beliefs, values, law, theory application and instrumentation which are shared by a given community within a particular activity".

A paradigm is a world view which researchers take for granted and is not revealed to the public unless certain circumstances require it. In this paper I have adopted two approaches for the revelation of paradigms. Firstly, by selecting public statements of researchers and secondly, by analysing research methodologies.

Recently an influential paper by Parlett and Hamilton has identified two paradigms in educational research; the "Agricultural/Botany" and the "Illuminative". At the risk of confusing readers, but with good reason, I choose to name the paradigms "Dominant" and "Innovative" respectively. It should also be noted that my comments apply equally to all disciplines used by educational research and to curriculum evaluation whom some would argue as being a distinct activity. This paper is about the assumptions that knowledge rests upon, not about specialities.

THE "DOMINANT PARADIGM".

For the past three decades research has been dominated by one paradigm which expresses itself in the work of its adherents by their use of the scientific method. Historically its dominance, and the parallel growth of educational research, can be attributed to a compatibility between the world view of political and educational decision makers since the war and the world view of researchers. A productive phenomenon which Gouldner would call "resonance". As we shall see there is a congruency of views between the 1945 Labour government policy of planning and the inherent characteristics of the dominant paradigm. Clarke, a founding father of educational research and a firm advocator of the "scientific approach", wrote in 1943.

"If we conducted our medical and engineering services and our industrial production with the same slipshod carelessness, the same disregard for precision of thought, the wild and reckless play of sentimentality or class prejudice or material interest masquerading as principle, with which we carry out our public discussions about education, most patients would die, most bridges would fall down and most manufacturing concerns would go bankrupt".

More recently similar sentiments are expressed by Yates, Director of the N.F.E.R., who sees scientific educational research generating knowledge to resolve educational problems. Under conditions where glimpses of the Dominant paradigm are revealed, a statement on the role of the N.F.E.R., there is the following:

"The National Foundation for Educational Research, is a research institution concerned with the study of problems arising at all levels within the national educational system of England and Wales. The Foundation seeks to promote the scientific study in a variety of ways including the development, maintenance and extension of a major research programme of its own, giving priority to problems of national importance. Projects are also undertaken for Government Departments and other international organisations".

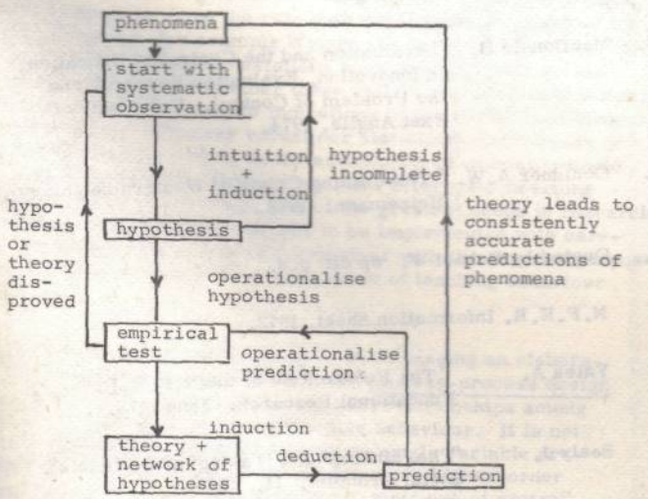
and

"A more realistic formulation would define our role as that of contributing to the evidence available to those whose responsibility it is to frame policy and devise procedures for its implementation. The findings of empirical research can inform their judgement but not provide a substitute for it".

These statements might not appear to be revealing, they have the element of commonsense. However, this statement of the role of the N.F.E.R., combined with an examination of what Yates means by scientific, is sufficient to reveal the Dominant paradigm.

What constitutes the scientific method is elaborated in Book E341 of the Open University's course on Methods of Educational Enquiry. Again, this is a condition where the paradigm is revealed to "outsiders". (However, in this case the outsiders are being initiated and so will be brought into the paradigm). The scientific method is represented in diagrammatic form and reproduced below. I am not attempting to establish some factitious relationship between unconnected works. Firstly, members of the N.F.E.R. act as advisers to the Open University's course team - thus there is a direct connection through personnel. Secondly, members within the same paradigm can be identified by their incestuous cross referencing and referal to definitive literature.

Fig.I: THE HYPOTHETICO-DEDUCTIVE METHOD



As we can see there are a number of basic stages of observation through what Popper calls the 'private act" of intuition and induction, to the construction of a hypothesis which Popper insists must be capable of refutation, to its operationalisation and testing. Confirmation of a hypothesis leads to the development of a network and subsequent theory on observed phenomena.

For both Yates and the Open University the social world, of which education is a part, is prone to problems. These appear for researchers within the paradigm to be "out there" as naturally occuring "facts", problems or phenomena, created either as part of the educational system or some aberration in it. All these are taken as given to the researcher.

This assumption that facts can be defined from a credible perspective or that a definition can be achieved through consensus is an essential and central one. It initially determines the relationship between educational research, society and the educational system. It is also necessary to the procedure of the scientific method, facilitating the construction of a hypothesis.

For the N.F.E.R. in particular this is an imperative stance in order to preserve its position of political neutrality. Consequently, and as a stated part of its policy, the N.F.E.R. is a "problem taking" rather than "problem making" institution. For the educational scientist this ensures the values of his method of objectivity and freedom from value bias. In the concluding sections of this paper I will return to these problems which have long been recognised. Wilson and Myrdal both offer solutions; the first author urges that the researchers should declare their value bias and by doing so somehow ostracize the ghost of subjectivity. The second accepts that value judgements are inevitable but recognising and accepting them is an honest position.

Unfortunately, this methodology reduces the social and psychological world to a world from which it it is borrowed - a world of "things" of the physico-chemical sciences. This places researchers in a schizoid and ambivalent position. Having reduced the world to things, with causal explanations of change brought about by active independent variables operating on passive dependent variables, researchers place themselves in the world only at the risk of being reduced to a "thing". Admittedly researchers are protected by their initial assumption of detachment but as people they have to make the further assumption that their "scientific" knowledge bestows them with rationality and humanity.

The Dominant Paradigm is, therefore, essentially conservative and incompatible with any other theory that might hold up the educational world as being one of multirealities. A world of multirealities is just what the "Innovative" Paradigm seeks.

THE INNOVATIVE PARADIGM

Members of the Innovative paradigm hope to circumvent criticisms made of the dominant one - but there is a paradox in its examination. As a new paradigm for educational research its members have articulated it publically and its assumptions, beliefs and values are more readily available. As a method it is characterised by its lack of form. Parlett and Hamilton do set out a number of stages for their strategy of curriculum evaluation but on their own admission these become blurred and are subject to backtracking and revision. I will then state briefly the features of the Innovative paradigm.

The researcher enters the field uncluttered by 'a priori' assumptions that characterise the Dominant paradigm about absolute reality. The investigation is open ended assuming that within a pluralistic educational system the possibility that different groups and individuals define situations differently. These multiple realities are collected to build

an over all comprehensive picture. From this position researchers are able to move to more and structured mode of investigation. Eclectic use is made of various data gathering instruments including conventional achievement and attitude test ".. but they have no privileged position, they merely represent another source of data".

The Innovative paradigm puts forward a view that reality is subjective and as this is dependent upon individuals all realities are relative, but with equal standing to one another.

As I indicated the Dominant paradigm reflected the world view of political and educational decision makers; so the Innovative paradigm reflects the humanism of the social sciences that came to prominence in the nineteen sixties. A period where there was a mass rejection of the Western world's scientific utilitarian culture expressed by student revolt, psychedelia and dropping out. The historical conditions brought about the situation as Martin puts it:

> "Positivism, objectivity and reason are still widely suspect. Theories and philosophies stressing subjectivity and relativism are now entirely respectable. In sociology, to give only one example, phenomenology and ethnomethodology may have had a long history in German philosophical thought, but it took the university troubles of the 196O's to raise them alongside the accepted classics as a normal part of the curriculum".

TOWARDS A "THIRD" PARADIGM.

The brevity in which I have outlined the "I nnovative" paradigm is intentional. In this concluding section I want to draw the readers attention to the similarities between the two paradigms; brevity is necessary to avoid replication. An account of the similarities between paradigms also explains the question mark in the title of this paper.

Superficially, the Dominant and Innovative appear to be diametrically opposed values - objectivity v subjectivity, absolutism v relativism, determinism v free will, and so on. The scientific method following distinctive, progressive and inflexible stages, striving to make the educational predictable. The methodology of the Innovative paradigm, on the other hand, is diffused, flexible, frequently turning back on itself in an effort to understand an unpredictable world. The se dichotomies are accepted by educational researchers, and if there has been debate between advocates, of either paradigms it has been from a position of inflexible commitment.

Practitioners of the Dominant paradigm assume their research findings can be developed into a general theory of education. Failure so far to do this is acknowledged by the Open University course team. Never-the-less the achievement of a general theory remains the ideal and realisable when researchers can make the recalcitrant social world of education more amenable by improved instrumentation and powerful statistical analysis. The methodological stance is maintained in the belief that objective data can be collected by imposing a definition of a situation on the world. As I have already pointed out this accepts-takes-for-granted-given definitions.

Practitioners of the Innovative paradigm do not accept given definitions but elicit them from the subjects of their research. But they do take for granted the taken-for-granted definitions of their subjects. Parlett and Hamilton, however, write that their illuminative method is capable of producing generisable statements - it has the same quality as the scientific method. It is difficult to see how this can be achieved as, to do so would require a shift from subjective relativistic data, without collapsing the illuminative back into the scientific method. If this is the case, and all writings within Innovative paradigm to my knowledge do not contradict this conclusion, we are faced with not two paradigms but different facets of one. The achievement of the Innovative paradigm to date has been to "make public" Popper's "private act" of intuition and induction.

Whether this conclusion is accepted does not alter the fact that both paradigms are caught in what can be termed as the trap of methodological individualism. Both paradigms are asocial and ahistorical in their data collection and analysis of education accepting the world as it is. (Albeit an imposed definition on the part of the Dominant paradigm, the acceptance of definition of those in education by the Innovative). This focus upon education as a discrete and independent system in society - a failure to see education as a historical product in a social context - means that it can only be understood and explained in the final analysis, by referal to intrinsic characteristics of individuals. Such an uncritical analysis is patently false although it is compatible with a society where the quality of the individual and "individuality" is held to be a central value. It also reflects the prime position that psychology has in educational research.

The way out of the present position of educational research is for a radical change of methodology. I accept the Innovative paradigm only as a starting point, but the realities it uncovers in education must be related by two dimensions of what Lefebur calls the vertical component of history and the horizontal component of the social context. This means that all the features of education - its organisation, examination system, power relationships, descriptive concepts and so on - will have to be located in the historical and social network that creates, maintains and controls it. By this method only can a relationship be shown between the so called objective "facts" and individuals subjective understanding of them.

CONCLUSION

Many of the ideas contained in this paper will need elaboration which space or time does not permit. I anticipate that what I advocate as a method will be seen by many as not being educational research. I certainly do not conceive educational research as a distinct and separate activity of the social sciences. Organised education is a phenomenon of our society that needs to be understood and explained and this can only be done in the larger context. Failure to do this will make educational research an exploitive activity.

BIBLIOGRAPHICAL NOTES

1. Brehaut W. "British Research in Education. Some aspects of its Development in Educational Research in Britain". Vol. 3. Butcher H.J. and Pont H.B., University of London Press. 1973.

2. Kuhn T.S. "The Structure of Scientific Revolutions". University of Chicago Press. 197O.

   It is inappropriate to debate Kuhn's theory in this paper. I have used his definition for convenience but this paper implicitly refutes his theory.

3. Partlett M. & Hamilton D. Working Paper No. 9. University of Edinburgh.

4. MacDonald B. "Evaluation and the Control of Education" in Innovation, Evaluation Research and the Problem of Control. University of East Anglia. 1974.

5. Gouldner A.W. "The Coming Crisis of Western Sociology". Heinemann. 1972.

6. Quoted by Rehaut W. op cit p.6.

7. N.F.E.R. Information Sheet. 1972.

8. Yates A. "The Future Role of the N.F.E.R." Educational Research. June 1973.

9. Seely J. "The 'Making' and 'Taking' of Problems". Social Problem. 14, 1966.

1O. Wilson J. "Three Myths in Educational Research". Educational Research. November 1973.

11. Myrdal G. "Objectivity in Social Research".

12. See Gouldner A.W. op cit pp. 496-497.

13. op cit.

14. Martin in "Black Paper 4". 1975.

15. In what I find an amazing statement the O.U. admits the scientific method might be inappropriate for the study of social phenomena but this question is best left to the philosophers. See "The Nature of Educational Research" p. 18 para. 2.13.

16. op cit.

17. Goldman L. "La Pensee des Lumieres". Annales Vol. 22, 4, (1967), pp 752-779. Unpublished English translation Geoffrey Tesson, University of Oregon.

18. Sartre J.P. 'Search for a Method" p. 51. Vintage Books, Random House 1968.

SCHOOL EXPERIENCE IN TEACHER EDUCATION

PRACTICAL EXPERIENCE IN THE PROFESSIONAL TRAINING OF TEACHERS: AN OVER-VIEW OF SOME RESEARCH ISSUES

Edith Cope.

Limitations in current school-practice procedures have already been exposed by research. They include role-confusion experienced by lecturers, teachers and students, segmentation of the practical from the theoretical elements in the course, idiosyncratic and semi-articulated criteria of evaluation, prodigal expense of time and energy (including petrol!) which nevertheless fails to ensure adequate guidance in preparation and feedback on performance: in short, a wasteful surplus of classroom experience over the capacity to reflect on and learn from that experience.[1] Action on these issues is restricted not by lack of awareness of the problems but by organizational inertia, by the relative status of institutions and by the power-groupings within them. Yet paradoxically research has also shown that school-practice is a highly valued element in teacher training courses.[2] Students learn from interaction with pupils and from experiencing the school from a new perspective even though formal arrangements for intensifying that learning are demonstrably inefficient. So with a high level of client satisfaction already established and with problems already identified, further research on school-practice would seem to have very low priority. Since my function at this conference is presumably something more than valedictory, what justification can I offer for the continuation of enquiries in this area?

Being a student is not the same experience as being a teacher. I have stressed how opportunities are lost if participants operate on the fallacy of non-difference.[3] Nevertheless, there is essential overlap. The student-teacher's task is to learn and to help children to learn. This is also the teacher's task, though the teacher's personal learning is no longer formally structured. Because of this overlap, research into school-practice should lead to exploration of issues fundamental to the whole educational transaction. It should explore how children learn, how teachers foster or inhibit that learning, how teacher-competence is recognised by pupils, parents, professional colleagues and the teacher himself, what is the relation, if any, between attributed competence and pupil-progress? If we knew how to help students to foster children's learning we should know something of how to help teachers. Logically such research should start with teachers and work back into the training system. The training system has, however, produced many of the researchers and may provide more feasible conditions for enquiry, so from expediency we have frequently worked from that end. School-practice as a point of inter-action where puils, students, lecturers and teachers engage in an inter-dependent educational process is too useful a locus to be abandoned, provided the emphasis is on those fundamental processes which are relevant to the total system. It is to stress this emphasis on the total system, and to cut across pre-service and in-service barriers, that this symposium includes not only Roy Griffiths' paper on developing integrated teaching practice programmes, but Ray Bolam's work on probationer induction.

But you will say, it is exactly these fundamental processes that are resistant to research. We all know Biddle and Ellena's notorious conclusion, on surveying over 2,000 studies, that we have no reliable measures of teacher effectiveness,[4] and ten years later we are no nearer. Exactly! Any search for "objective" measures of general applicability is the pursuit of a chimera. Valid criteria of effectiveness must, of course, be striven for in every educational undertaking. It is probable that in relation to teacher-competence measures of children's learning, as advocated persuasively by Brian Start and Ed Stones, if sufficiently context-bound, sensitive, and embracing the affective plus the cognitive, would take us some way. But the notions of effectiveness as surveyed by Biddle and Ellena are linked to a positivist and empirical concept of research which, while it should not be entirely supplanted, should be and is being supplemented.

The subjective element in scientific ways of knowing has been impressively expounded by Polanyi.[5] The built-in assumptions behind so-called objective data-gathering - "measurement by fiat" - have been exposed by Cicourel and the ethno-methodologists.[6] Statistical correlations may convey a largely spurious objectivity: more important, they may be irrelevant to the problems needing exploration. Job analyses through records, diaries and ethological observations have shown how limited a part of the teacher's day is spent in directly interacting with pupils in a cognitive learning situation. The school is a social world, and learning occurs in the context of human relationships. We need to develop research models appropriate to the complexities of this social world of school and classroom. Fortunately, the current state of flux, or even crisis, in sociology and social psychology not only allows for this, but positively demands innovative daring in place of respectability or mere inegenuity.

I agree with the view expressed by Roger Dale in his article in the Eductional Review[7] 1973, that a phenomenological perspective offers much promise. Such a perspective acknowledges that the social world is a construct sustained as well as experienced by its participant members.[8] Since humans have the unique capacity for language (I use the adjective advisedly in spite of Washo) and, through language, the capacity for intention, reflection and the attribution of meaning, it follows that in order to understand this social world, we need to acquire the versions of reality offered by its members. This means embarking on research which deals not only with behaviour but with purposes, interpretations, reasons.

"Similar behaviour has no necessary implications for similarity of reasons for that behaviour". (Harre, 1972).[9]

It necessitates utilising as evidence individuals' accounts of social reality as they perceive it. One schema for dealing with such accounts within a dramaturgical interactionist framework is offered by Harre and Secord.[10] They classify episodes as "formal" or "enigmatic", and within these categories distinguish rituals, routines, games and entertainments. Harre advocates that the social scientist should look for the equivalent of "scripts, liturgies, books of rules, and not for correlations, statistics and significant measures".[11]

Perhaps one of the problems for teachers at the moment is that lessons have ceased to be "formal" episodes with established scripts and rituals and have become increasingly "enigmatic", with participants uncertain of their lines or even insisting on enacting different improvisations! When this occurs, the consequences are at best embarassing, at worst traumatic. It is not only the individual psyche that is threatened, but the precarious stability of the whole social world.

The phenomenological perspective enables the researcher to see the participants of the social world of the classroom as the sustainers of that social world, creating it as well as being acted upon by it. This perception obliges the researcher to involve teachers on a very different basis from that which treats them as biro-wielding fillers-in of pre-set questionnaires and check lists. Since I am convinced that one of the most powerful ways of continuing teachers' education is to involve them both in student-training and research undertakings on a basis of active partnership, then a perspective which offers a validation of this belief is obviously attractive to me. Majority participants in the sustaining of the social world of school and classroom are, of course, the pupils, and a recognition of this means that their versions of the reality of the social world must also be taken seriously by researchers. I acknowledge the delicacy of this, but believe we have reached a stage where we should shift some attention to the majority of the cast!

All this may seem threatening when compared with the lost protective pretence of researcher objectivity, and with the assurances offered by statistical matrices. It can, however, be experienced as liberating. Yet in no sense am I advocating abandoning the striving for objectivity, for generalisable statements, for descriptions and explanations which improve on those offered by participants, for the utilisation of measures where appropriate Phenomenology has severe limitations of scope as succinctly expressed by Dale in the same article.

> "Phenomonological sociologists concentrate on the surface structure rather than the deep structure of social life. They emphasise the immediate experiences and subjective meanings of everyday life; they do not seek to reveal the nature of the system of social relationships which governs the way those meanings are distributed within societies". [12]

What I am suggesting is that new perspectives provide a correction for previous astigmation and myopeia, but, as the vision changes, they themselves become a source of distortion. The only permanently valid stance is still contained in C. Wright Mill's Maxim, "Let every man be his own methodologist; let every man be his own theorist" [13] since this imposes on the researcher the responsibility for choosing, or creating, a theoretical stance, explanatory models and a set of procedures appropriate to the specific problems being studied. This may well include utilising evidence gained by both subjective and more objective techniques. One particular instance which I can develop in talking to this paper arises out of an action research project which I initiated at Bristol but which was carried through by Ian Lewis, now back at York. In this, small groups of teachers, students and lecturers structured sequences of work on the development of reading skills in junior school children, and planned the accompanying supervisory procedures which were to be the main focus of the enquiry. Evaluation took the form of individual unstructured interviews in which the adult participants' views were tape-recorded. But in addition tape-recordings were made of supervisory sessions in the workshop situation. A sample from these was analysed utilising a technique developed by Richard Weller, now Associate Professor at the University of North Carolina. His system is a development of interaction analysis, but applied not to classroom exchanges but to the supervisory conference. He calls the system M.O.S.A.I.C.S. - Multi Dimensional Observation System for the Analysis of Interaction in Clinical Supervision, and it is fully explained in his book Verbal Communication in Instructional Supervision, 1971.[14] Ian Lewis' utilisation of this aspect of the evaluation of the Bristol project exemplifies that open-minded combination of more subjective techniques with more objective, structured evidence which I believe we need. It is this willingness to implement diverse models which is essential. Otherwise we reject, through ideology, approaches which may yield insights.

I should like to end, however, with two assertions. The first is that we must be willing to implement research which not only implies values, but explores and asserts them. Fortunately the climate is increasingly sympathetic. Asplund, writing on the moral element in the work of Sorel, Durkheim, Pareto and Weber, once a source of embarrassment and even scandal to twentieth century social scientists, comments on the renewed interest in these elements and goes on to affirm that "a value-relevant social science may be both possible and legitimate". [15]

The second is that since, as educationalists, we wish not only to understand the world but to change it, then action originating from a phenomenological stance provides us with the only valid way to promote innovation which "takes". Change in the social world occurs and is maintained only when participants co-operate to sustain an alternative version of reality. The researcher can be the catalyst who not only assists participants to a self-conscious awareness of the nature of the realities they are experiencing, but collaborates with them in devising new performances.

REFERENCES

(1) For supporting evidence and further discussion, see:-

COPE E — School Experience in Teacher Education University of Bristol, 1971

COPE E — A Study of a School-Supervised Practice University of Bristol, 1971

STONES E and MORRIS S — Teaching Practice: Problems and Perspectives Methuen, 1972

STONES E — How Long is a Piece of String? Paper presented at S.R.H.E. conference on Research into Teaching Practice, December, 1974 (To be published)

(2) COPE E — School Experience in Teacher Education p. 6 and p.106

(3) COPE E — ibid, pp. 40-41

(4) BIDDLE B J and ELLENA W J — Contemporary Research on Teacher Effectiveness Holt, Rinehart and Winston, 1964

(5) POLANYI Michael — Personal Knowledge. Towards a Post-Critical Philosophy Routledge, 1958

(6) CICOUREL A — Method and Measurement in Sociology Free Press, 1964

(7) DALE J R — Phenomenological Perspectives and the Sociology of the School Educational Review, Vol. 25, No. 3, 1973

(8) For a spelling out of the implications of this stance, see:-

FILMER Paul; PHILLIPSON Michael; SILVERMAN David; WALSH David — New Directions in Sociological Theory Collier-Macmillan, 1972

(9) HARRE R — The analysis of episodes in Joachim Israel and Henri Tajfel (Eds.) The Context of Social Psychology: A Critical Assessment Academic Press, 1972

(10) HARRE R and SECORD P F — The Exploration of Social Behaviour Blackwell, 1972

(11) HARRE R — The analysis of episodes, op. cit. p.422

(12) DALE LR — op. cit. p. 186

(13) MILLS C WRIGHT — The Sociological Imagination Oxford University Press, 1959, p224

(14) WELLER Richard — Verbal Communication in Instructional Supervision Teachers College, 1971

(15) ASPLUND J — On the concept of value relevance in J Israel and M Tajfel, op. cit. p.261

AN ANALYTIC APPROACH TO AN INTEGRATED PRACTICE TEACHING PROGRAMME

Roy Griffiths.
University of Stirling.

A substantial component of all British teacher training courses comprises observation and teaching in schools. The importance and value of such experience seems widely accepted by teacher organisations, teacher trainers and various working parties. Moreover, day-to-day experience and research results testify to student teachers themselves placing a high value on school practice.

Yet, as Morrison and McIntyre (1973) point out, the importance which is attached to school practice should not be taken to imply satisfaction. They claim that respondents have been as critical of the content, procedures and arrangements of school practice as they have been of other aspects of college curricula. Cope (1975) notes in her contribution to this symposium that research has already exposed a variety of limitations in current school practice procedures. Her own research at Bristol (Cope, 1971a, 1971b) has done much to document these, particularly those arising from the problems in college-school liaison Hooper and Johnston (1973) suggest that school practice functions mainly as an implement of social control and a means to learning professional conformity. They advocate 'a radical reappraisal of that which has long been regarded as inviolate'. Lewis (1975) has argued that,

> "As a learning situation it is clear that teaching practice signally fails to meet even the most rudiamentary set of objectives.....for the student the operative word in the title is 'teaching' rather than 'practice', since there has rarely been any opportunity to acquire any of the skills which might be practised".

Because of the variety and complexity of the deficiencies in school practice, it is seemingly impossible to posit a single solution. Rather, different deficiencies have to be attacked on different fronts. Lewis (1974), for example, concentrates on the lack of working co-operation between schools and colleges, and reports the study he carried out at Bristol to develop ways of organising a range of opportunities for extensive and co-operative work involving teachers, lecturers and students. Other papers (for example, Poppleton 1968, Stones and Morris 1972) have concentrated on the problems of assessment systems. Developing an alternative to the traditional college supervisor role was the focus of the study by Caspari and Eggleston (1965), and Cope (1973) surveys various approaches to assisting learning which are now available.

The present paper proposes an attack on some of the deficiencies of school practice on another front. It is based upon moving away from reliance on the school classroom as the only training ground. It seems unnecessary to restrict all of those activities which can be labelled 'practice teaching' to school experience. For example, Brown (1975) shows how many aspects of the skills of performance, perception and preparation involved in teaching can be practised in college settings. Such practical work experiences should be viewed not as replacing school practice, but as preparatory to and supplementary to school based work. To illustrate, Travers (1975) argues for attention in pre-service work to the development of teacher role characteristics, and describes a series of activities only after which "the student of education should have reached a level at which the student can profit from work in the classroom". As Peck and Tucker (1973) conclude from their research review, "Teacher education can no longer remain in a happily ignorant, ineffectual state consisting of romanticized lectures on the one hand, and fuzzy or uplanned 'practical' experience on the other".

The aim of the approach I am discussing is, the development of a variety of interrelated practice experiences (including school practice) which could be synthesized into an integrated practice teaching programme as appropriate in particular circumstances. Thus the attempt is not to prescribe a blue print programme, but to identify elements from which particular situation specific integrations may be created.

In order to minimize the tendency towards fuzzy or unplanned experience I have suggested in my title an analytic approach. This would involve procedures such as the analysis of practical experience in terms of its objectives, carefully planned training procedures aimed explicitly at these objectives, assessment of the results of the work, and provision of feedback to the learners.

Some of the techniques and approaches upon which such a programme of practice teaching could be based include:-

1. Microteaching and various adaptations such as mini-teaching and 'bridging' experiences (Brown 1975a and 1975b, Walker and Adelman 1975).
2. Training in decision making including lesson planning (Hill and Martin 1971, Waimon et al 1972, Bowles 1973).
3. Protocol materials and cognitive discrimination training (Burdin and Cruickshank 1974, Borg and Stone 1974, Borg 1974, Wagner 1973).
4. Development of conscious self-monitoring (Elliott and Labbett 1975).
5. Training in analyticity (Cameron-Jones 1974).

\- and of course, school practice.

The time available for talking to this paper in the symposium will be used to elaborate on some of the techniques and approaches listed above, and to outline some areas of research which could facilitate the development of integrated practice teaching programmes.

REFERENCES

BORG W R (1974) Field Test Report 1972-73 Protocol Materials Project Utah State University

BORG W R and STONE D R (1974) Protocol materials as a tool for changing teacher behavior Journal of Experimental Education, 43, 1 34-39.

BOWLES F D (1973) Decision making in instruction Journal of Teacher Education, 24, 1, 38-40

BROWN G (1975)a Microteaching - A Programme of Teaching Skills Methuen

BROWN G (1975)b Microteaching: research and developments, in Chanan G and Delamont S Frontiers of Classroom Research, N.F.E.R.

BURDIN J L and CRUICKSHANK D R (1974) Protocol materials: training materials for uniting theory and practice ERIC, ED 093 812 several contributions from which appear in the Journal of Teacher Education, 1974, 25, 4

CAMERON-JONES M (1974) Teaching Teaching in Gibbs I, Conference on Microteaching/ Interaction Analysis Callendar Park College of Education

CASPARI I E and EGGLESTON J (1965) A new approach to supervision of teaching practice, Education for Teaching, 68, 42-52

COPE E (1971a) School Experience in Teacher Education University of Bristol, School of Education

COPE E (1971b) A Study of School Supervised Practice University of Bristol, School of Education

COPE E (1973) School experience and student learning, in Lomax D.E. The Education of Teachers in Britain Wiley p. 243-261

COPE E (1975) Practical experience in the professional training of teachers: an overview of some research issues Paper presented at the B.E.R.A. Annual Conference, University of Stirling

ELLIOTT J and LABBETT B (1975) Teaching, research, and teacher education: some comments on competency based teacher education Education for Teaching, 96, 51-61

HILL J and MARTIN W R (1971) Training for educational decision making Journal of Teacher Education, 22, 4, 443-447

HOOPER D and JOHNSTON T (1973) Teaching practice: training or social control Education for Teaching, 92, 25-30

LEWIS I (1974) The role of school practice in the training of teachers Education for Development, 3, 2, 3-17

LEWIS I (1975) Teacher training: professional or peripheral? Education for Teaching, 96, 35-42

MORRISON A and MCINTYRE D (1973) Teachers and Teaching 2nd Edition, Penguin

PECK R F and TUCKER J A (1973) Research on teacher education, in Travers R M W, Second Handbook of Research on Teaching. Rand McNally p. 940-978

POPPLETON P (1968) The assessment of teaching practice: what criteria do we use? Education for Teaching, 75, 58-64

STONES E and MORRIS S (1972) The assessment of practical teaching: Educational Research, 14, 2, 110-119

TRAVERS R M W (1975) Empirically based teacher education: The Educational Forum, 39, 4, 417-433

WAGNER A C (1973) Changing teaching behaviour: a comparison of microteaching and cognitive discrimination training Journal of Educational Psychology 64, 3, 299-305

WAIMON M D, BELL D D and RAMSEYER GC (1972) The effects of competency-based training on the performance of prospective teachers Journal of Teacher Education, 23, 2, 237-245

WALKER R and ADELMAN C (1975) A Guide to Classroom Observation Methuen

THE SUPERVISORY ROLE OF THE TEACHER TUTOR – A COMPLEX INNOVATION

R. Bolam.
University of Bristol.

1. INTRODUCTION: THE CONTEXT

My purpose in this paper is to give an outline account of one aspect of the Teacher Induction Pilot Schemes (T.I.P.S.) Project - the introduction of the teacher tutor role. Before doing so, it will probably be helpful to describe briefly the aims and scope of the overall project. The T.I.P.S. Project is funded by the D.E.S. and based at the Research Unit in the University of Bristol, School of Education. The brief of the Bristol team, which consists of myself and my colleague, Keith Baker, is to promote, co-ordinate and inform the monitoring and dissemination activities of the two 'official' pilot schemes in Liverpool and Northumberland and of ten 'unofficial' L.E.A. schemes around the country. In so doing, we work closely with the local evaluators in Liverpool (David Hill) and Northumberland (Colin McCabe) and also with the programme organisers in all twelve L.E.A.s.

From the Bristol end, we have seen the project as being essentially concerned with the introduction of a massive and complex innovation which in itself contains several complex sub-innovations, each working to separate timetables. Accordingly, we are attempting to monitor the overall project's activities at six principal levels: national, regional, local authority, professional centre, school and individual probationer (cf. Bolam, 1975). This paper, therefore, deals with only one, limited aspect of the overall project - the attempt to introduce the innovative teacher tutor role into the schools involved in the two 'official' pilot schemes.

2. THE RESEARCH DESIGN: THEORETICAL AND METHODOLOGICAL CONSIDERATIONS

The project has been conceptualised as an exercise in the management of the implementation of a complex innovation.

The problems of managing the change process may be usefully explored via the following two-dimensional framework. In any innovation process we can distinguish between four major factors: the innovation; the change agent; the target user and the innovation process over time. The first three factors may be defined as open systems with internal and external characteristics. These factors interact with, and are changed by, each other during the course of the innovation process over time (see Bolam, forthcoming).

Figure 1 A Conceptual Framework for the Study of Educational Innovation

		Dimension 1: The Three Major Systems		
		System 1	System 2	System 3
	Time 1: Before The Antecedent Stage	The Innovation	The Change Agent	The Target User
Dimension 2: The Process of Innovation Over Time	Time 2: The Inter-Active Stage	The Innovation	The Change Agent	The Target User
	Time 3: After The Consequent Stage	The Innovation	The Change Agent	The Target User

In essence, the framework generates four sets of questions about the innovation, the change agent and the target user:-

1) What are their most salient and significant characteristics in the context of the particular innovation process under consideration?

2) What are they like prior to the start of the process?

3) What happens when they interact with each other during the process?

4) What are they like after that interaction?

The framework is being used as the basis for the analysis of each of the six levels of the overall project but in this paper it is applied only to one level and one sub-innovation. The sub-innovation is defined as the teacher tutor role model, the change agent system as the L.E.A. programme organiser and the target user system as the teacher tutors in the two L.E.A.s. Spacedoes not permit a complete account of an application of the framework to this particular sub-innovation, so only certain key issues will be highlighted. A useful source for these key issues is provided by Gross et al. 's, 1971, study of the introduction of a new role for teachers in an American elementary school. The study concludes that the innovation failed and of the reasons which it offers, the following are relevant for our present purposes:-

1) the teachers did not have a clear understanding of what was expected of them in their new role;

2) in any case, they did not have the necessary skills to carry out their new role;

3) the organisational arrangements in the school (e.g. the inflexible timetable) were incompatible with the innovation;

4) As a result of these unsatisfactory experiences, resistance to the change developed amongst the staff.

Gross et al. argue that it is a management task to plan, support and monitor the innovation process and, therefore, that it is a management responsibility to clarify the teachers' role definition, to provide access to training for the new role, to modify the organisational arrangements and, finally, to devise feedback mechanisms to identify these and other difficulties as they arise, thus minimising the possibility of resistance developing during the implementation process. They conclude that the task of management is of critical importance in all of these activities.

A second source of likely major issues directly relevant to the particular innovation under study was provided by work on teacher induction and training. Previous research carried out at Bristol on the induction year (Taylor and Dale, 1971; Bolam, 1973) led to the conclusion that external induction support, though an essential component of an overall programme, is not in itself sufficient. External tutors and centre-based courses cannot easily provide the specific and individualised help which probationers seem to need. We may hypothesise, therefore, that the particular strength of the school-based teacher tutor is his ability to provide individualised help which is informed by an intimate knowledge of a specific school and classroom context. This hypothesis is one which accords well with recent work in the U.S.A. on clinical supervision (e.g. Goldhammer,1969; Mosher and Purpel, 1972; Cogan, 1973) which offers a useful framework of practice and concepts. The essentials of clinical supervision may be summarised in the following quotations from Mosher and Purpel, 1972:-

1) Its 'primary objective ... is the improvement of instruction' (p.78)

2) It 'focusses on what and how teachers teach as they teach' (p. 78)

3) 'It is this principle of direct application that makes the method clinical: it addresses the doing or practice dimension of teaching' (p.79)

4) 'adherents believe that the analysis of teaching can be rigorous and systematic, that it should be ongoing, that it requires specific analytical skills ...' (p. 79)

5) It consists characteristically of a three-stage ongoing 'cycle of systematic planning, observation and critical analysis of teaching' (p. 81)

The central focus of this paper is upon the attempt to implement an innovative and complex role - the supervisory role of the teacher tutor - into two local education authorities. The theoretical framework outlined above is used as the basis for a limited exploration of some key barriers and problems likely to be encountered in the innovation process, notably those suggested by Gross et al., 1971, in their study of an analogous innovation, by previous Bristol research on induction and by American work on clinical supervision.

It will by now be clear that the project is essentially concerned with action research. The theoretical, methodological, practical and political problems of the action researcher are well known and I will not rehearse them here. With respect to the project as a whole, it has been agreed with the local evaluators in Liverpool and Northumberland that the differences between the various schemes are such that controlled experimental studies will probably not be possible either within or between L.E.A.s. The monitoring design is, therefore, based upon separate case-studies of each L.E.A.s scheme using an evaluative research approach (Weiss, 1972). The task of monitoring the nature and quality of the interaction between teacher tutors and probationers obviously poses particular methodological problems. Given our other monitoring commitments, we simply do not have the resources to carry out systematic observations of even a sample of tutors and this would, in any case, be a questionable procedure. We have up to now concentrated on collecting data via interviews, questionnaires and diaries. The rest of this paper draws upon such data and also upon our experience of video-taping some tutors in real and simulated interchanges with probationers.

3. IMPLEMENTING THE INNOVATIVE ROLE

The first job facing the L.E.A. programme organisers was to define the teacher tutor role and set up a training programme: that much was clear from the D.E.S. brief. However, the James Report was vague and the subsequent White Paper said nothing about the details of the tutor's induction role. Consequently, most of the conferences and A.T.O./D.E.S. Courses which followed were exploratory in their attempts to define the role. Thus, when the Liverpool and Northumberland Advisory Committees for the pilot schemes set about their task of advising the programme organisers, they really had very little to guide them. Inevitably, much of their early thinking and discussion dealt with such matters as the appointment and payment of tutors and relatively little detailed advice was provided for the 400 or so tutors in the two areas. However, the Liverpool Advisory Committee did suggest that they should assist 'the new teacher, either directly or through specialist colleagues, to develop effective teaching techniques in the classroom' (Liverpool Education Committee, 1975, p. 11) and it also agreed that tutors should be familiar with:-

> 'Formal and informal techniques for the analysis of teaching and learning situations and for the subsequent diagnosis of individual probationers' strengths and weaknesses'.

and

> 'Techniques of counselling and group leadership.'

The Northumberland guidelines were roughly similar and, understandably, the job of filling in the details of this outline in both schemes was left to the training course organisers, and, ultimately, to the tutors themselves when they actually came to carry out their new roles. How, then, did this work out in practice?

First, let us consider briefly the training courses which were organised for the tutors. It is worth recalling the climate of opinion in the profession at the time. On the one hand, the National Union of Teachers, 1973, argued that courses for tutors should last for at least one month. On the other hand, the National Association of Schoolmasters, 1973, was doubtful about the need for any training at all. In the event, Liverpool organised a 10-day 'briefing' programme and Northumberland ran a 5-day 'initial preparation' programme. Although courses differed both between and within (e.g. for primary and secondary tutors) the two schemes, one striking feature was common to all courses: only a small proportion of the total time was given over to training or practice in what we may call supervision skills. The vast majority of the Course Sessions dealt with exploratory discussions of the tutor's role, lectures on such topics as probationers' needs and external support agencies (e.g. the Welfare Service) and visits to colleges of education.

What of the way tutors actually interpreted their role in action?

The clinical supervision framework is not directly applicable but it does indicate certain key questions and issues which we may regard as indices of the extent to which tutors favour and implement the supervisory role. The Liverpool evaluator reports (Hill, 1974) that over two thirds of his sample of 178 tutors saw their most important function as being 'to counsel, encourage, advise and reassure' the probationers. About one fifth said it was 'helping with organisation, discipline, lesson planning, and generally fostering management skills in the classroom', but not one ticked the following items:-

> 'Observing and commenting upon the amount of learning taking place in the probationer's classroom.'
>
> 'Joining in with the probationer in his room in a teaching situation.'
>
> 'Arranging for the probationer to observe me teaching.'

Only four tutors saw their main job as:-

> 'Offering advice and guidance with teaching techniques.'

In Northumberland, the local evaluator discovered (McCabe, 1975) that probationers generally disagreed with the statement:-

> 'Someone should watch you teach and discuss your performance with you.'

although high school probationers were less inclined to disagree than first and middle school probationers.

Intensive pilot interviews of over half the staff in seven schools provided data which largely corroborates these questionnaire findings. For example, at the end of the school year we interviewed staff in one primary school in which the teacher tutor was clearly both extremely conscientious and competent. She had not, however, either observed her probationer teach or arranged for her to observe a colleague teaching. But she did stress that she helped her probationer to evaluate her own performance during their regular and systematic discussion session each Monday morning.

> 'I have no set scheme, but I do try to help her to be clear about her aims, how far she achieves them and what, if anything, has gone wrong.'

The tutor pointed out that, while she could, in principle, observe her probationer teaching, '... one does not want to intrude'. The probationer's view was that having the tutor observe her would be '... like going back on teaching practice', although she did say that she wouldn't mind regular visits.

The tutor's views on arranging for her probationer to observe either herself or another senior colleague teaching were:-

> 'If she is having difficulties it may be necessary but it could cause her to lose standing with her own class. If nobody else does, she is seen to be different.'

Another interviewee, this time an experienced teacher, considered that 'teaching is a personal thing and it is not desirable for a young teacher to copy an experienced colleague'. The oldest of our respondents in this school was a man who was very supportive of the overall scheme; he thought that both sorts of observation were best 'left out' because of the difficulties they could cause.

4. DISCUSSION

Before considering the substantive issues in more detail it is worth noting some of the features of this particular innovation process in relation to the theoretical framework outlined in Section 2 above. The systemic characteristics of both the change agent and target user are evident: the programme organisers (the change agents) have, for example, to take account of the advice of their advisory committees which in turn reflect wider professional views; the teacher tutors (the target users) each have to take into account the behaviour and opinions of their probationers and their school colleagues. Similarly, the way in which the innovation is differentially defined is evident: the teacher tutor role is apparently defined in one way by the Advisory Committee and in another by the tutors themselves.

This brings us to the central issue: although the teacher tutor role was defined by the programme organisers in a certain way it is doubtful whether the role was actually implemented in this way. In considering this issue, we turn the first of Gross et al.'s conclusions to answer the question - 'How clear were the tutors about what their role involved?' We may usefully distinguish between two major components in the teacher tutor role by placing them at opposite ends of a continuum:-

PASTORAL _____ TRAINING

The pastoral component involves helping the probationer with personal (e.g. accommodation) and professional (e.g. information about school rules) tasks and problems. The training component involves helping probationers to improve their general pedagogical skills (e.g. classroom management) and specialist subject skills (e.g. reading or physics). There appears to be fairly general agreement in the profession as a whole and in the pilot areas that the pastoral component is a legitimate part of the teacher tutor role. Moreover, teacher tutors appear to be emphasising this aspect of the role (although some are reluctant to carry out too many personal support activities) and frequently mention that they want to be first and foremost a friendly and supportive colleague. Hill's findings, above, support this view. The training component, which many regard as the raison d'etre of the tutors' role (and which is closely similar to the clinical supervision concept described earlier) was defined as an explicit part of the role in the pilot areas. But, although the tutors accepted that training was part of their role, they have tended to interpret it as involving discussions outside the classroom and they have apparently been disinclined to employ classroom observation as the basis for it.

We, therefore, appear to have a situation in which tutors are able to interpret their role rather than being unclear about it. This was in part due to the fact that, in a pilot scheme some flexibility was desirable anyway. But the reluctance of the Advisory Committees and of the teacher tutors to grasp the training nettle too explicitly probably also relates to certain strongly held values in the profession. The probationers were no doubt reluctant to be observed by their tutors for several reasons of a personal nature but the response from the interviewee above is a significant one. The most direct association of observation sessions in her mind was with her school practice experience as a student teacher. Thus, there is an implication echoed in the tone of the Liverpool Advisory Committees role specification, that classroom observation is inevitably linked with assessment. Furthermore, there is a clear suggestion, also echoed in the replies of the teacher tutor, that the probationer's professional status would be diminished, especially since none of the other staff either observed or was observed by colleagues. Obviously, practises vary: in some schools team teaching or an open plan organisation create a climate in which it is normal for teachers to work alongside and observe each other at work; in other schools the ultimate professional sin is to enter a classroom in which a colleague is teaching; in between there are a whole range of 'climates'. However, it has to be recognised that observation and the clinical supervision model will be interpreted by many teachers as a threat to their cherished classroom autonomy and is, therefore, likely to be unacceptable to many of them.

It is, therefore, not surprising that the briefing and preparation courses paid little attention to skill training for tutors. But there is also a set of reasons related to the difficulties of training in supervisory skills. The first reason is exemplified in the N.A.S. policy statement, referred to earlier, which reflects a fundamental disagreement that there is anything especially difficult about helping a probationer which a conventionally experienced colleague cannot handle. This is often coupled with a lack of awareness of such relatively recent developments as micro-teaching and interaction analysis. The second reason is that, so few people possess such skills anyway; it is hardly surprising that they did not figure prominently in the tutor training courses since the number of people in the country experienced enough to mount a sustained training course in, say, interaction analysis skills (i.e. as opposed to giving a descriptive lecture) is so small. In any case, although micro-teaching and interaction analysis ought to be components in a training course their

focus is too narrow for them to stand alone. The essence of the teacher tutor's task is that he has to deal with the whole of a probationer's experience in a real life context. He has to focus on the content as well as the methods used by the probationer. Thus, curriculum analysis skills are vitally necessary. The strength of clinical supervision is that it adopts this more comprehensive approach and recognises the practical constraints (e.g. the timetable) within which tutors have to work. Ideally, however, it too would be supplemented by micro-counselling training (Ivey and Rollin, 1974) and change agent training (Havelock and Havelock, 1973).

Finally, there are certain logistic and administrative factors which may have constrained the tutors from emphasising their training function. The timetable arrangements within schools frequently made it difficult for tutors to observe probationers teaching and vice-versa. In the interview school described above, for example, this was the case: The head there had arranged that tutor and probationer should be free together in order that they might hold a discussion session. This was, in itself, an excellent idea but it did make it difficult for the tutor to adopt a clinical supervision approach if she had been so inclined. Another administrative factor was the attention that had, of necessity, to be given to organising the centre-based courses for probationers. This was an enormously complex logistic operation which may well have had two unintended consequences: first, because it was so demanding in administrative terms and because a day conference or a series of course sessions for probationers provides immediately obvious and visible evidence that 'induction' is taking place, the external courses were in danger of becoming regarded as the major component of the pilot schemes; second, there is evidence that the tutors came to feel that their contributions in consequence, were being undervalued precisely at the time when many of them were uncertain as to how best to proceed with their new role.

The conclusions reached by Gross et al, 1971, are all relevant to an understanding of the implementation and management of the innovative teacher tutor role but they tend to oversimplify a highly complex process. It is far from certain that 'role clarity' is desirable in a pilot study since it may lead to premature rejection of an innovation whereas a flexible approach can lead to increased understanding of the professional issues involved. The need for training in skills, however desirable, can only be met if trainers are available to provide it. Here, too, it is probably better to move fairly pragmatically utilising whatever resources are available rather than to reject the innovation until an ideal situation is reached. The administrative constraints described above were unintended outcomes and were, arguably, extremely difficult to avoid if not predict. Finally, because the scheme is being internally monitored by the L.E.A. programme organisers they are aware of these issues and are taking steps to deal with them in Year 2 of the pilot scheme.

5. CONCLUSIONS

This has necessarily been a sketchy and incomplete account of one limited aspect of the T.I.P.S. Project: the implementation of an innovative teacher tutor role in the first year of the induction pilot schemes. A conceptual framework for the study of innovations and issues derived from the literature on innovation, induction and clinical supervision were used as the basis for the study. It was concluded that, although the programme organisers recommended that the role should include a significant training component, this has been interpreted by many tutors as not including classroom observation (or clinical supervision) activities and reasons were offered to explain this. It was also concluded that, although study by Gross et al, 1971, identified important innovation management issues, its conclusions tended to oversimplify the problems involved in a complex task.

Finally, and by way of an epilogue, it is perhaps worth considering briefly some underlying reasons for the problems associated with training teacher tutors. It is clear that there now exists a whole battery of techniques for training the trainers of teachers. A number of them are the subject of papers at this conference and several are now over ten years old. Yet when a major national pilot scheme is mounted we find that these techniques are hardly used at all. In seeking the reason for this state of affairs it may be helpful to regard these techniques as innovations in themselves and to analyse them in terms of some well known models of the process of knowledge diffusion. Havelock, 1971, distinguishes between three principal models:- social interaction, research development and diffusion and problem-solving user. At present, techniques for training trainers are still largely the preserve of researchers and their ideas and findings tend to diffuse through the system mainly by unplanned and informal social interaction networks. With one or two exceptions the development and diffusion stages of the R.D. and D. model have not been systematically tackled. As far as I know there has been no attempt in this country to adopt a problem-solving approach. In the induction schemes this would mean working with those involved to create training approaches which are viable and acceptable in the context of the real-life situations within which the tutors work. Thus, the attitudes of teachers to classroom observation and the constraints of time and skill-levels which are likely to obtain with average teacher tutors in schools, would both have to be taken in account. Similar questions would have to be faced with respect to initial and in-service training. In short, are modern analytic techniques likely to be generally used by teacher trainers? Or are they going to be used almost exclusively by researchers?

Bibliography

Bolam, R., 1973, Induction Programmes for Probationary Teachers. Bristol: University of Bristol, School of Education

Bolam, R., 1975, 'The Teacher Induction Pilot Schemes Project', London Educational Review 4.1

Bolam, R., (forthcoming), 'The study and management of educational innovation: towards a conceptual framework', in the Readers for Open University Courses E203 and E321

Cogan, M., 1973, Clinical Supervision, London: Houghton Mifflin Publishers Ltd

Goldhammer, K., 1969, Clinical Supervision. New York: Holt, Rinehart and Winston.

Gross, N., Giacquinta, J.B. and Bernstein, M., 1971 Implementing Organisational Innovations. London: Harper and Row

Havelock, R.G., 1971, 'The utilisation of educational research and development', British Journal of Educational Technology, 2.2

Havelock, R.G. and M.C., 1973, Training for Change Agents. Ann Arbor, Michigan: University of Michigan Institute for Social Research.

Hill, D., 1974, Evaluation Report Number 8: Perceptions of Teacher Tutors Towards Their Role. University of Liverpool, School of Education

Hoyle, E., 1973, 'Strategies of curriculum change': Watkins, R. (ed) In-service Training: Structure and Content. London: Ward Lock

Ivey, A.E. and Rollin S.A., 1974, 'Counselling technology: micro-counselling and systematic approaches to human relations training', British Journal of Educational Technology, 5.2 15-29

Liverpool Education Committee, 1975, Pilot Schemes for the Induction of New Teachers: Reports Numbers 1 to 6. Liverpool Education Authority

McCabe, C., 1975, Attitudes to Induction. Draft report. University of Newcastle, School of Education

Mosher, R. and Purpel, D., 1972, Supervision: The Reluctant Profession. London: Houghton Mifflin Publishers Ltd.

National Association of Schoolmasters, 1973, The Induction of New Teachers and the Role of the Professional Tutor

National Union of Teachers, 1973, The Induction of New Teachers

Taylor, J.K., and Dale, I.R., 1971, A survey of Teachers in Their First Year of Service. Bristol: University of Bristol, School of Education

Weiss, C.H., 1972, Evaluation Research - Methods for Assessing Program Effectiveness. Methods of Social Science Series. N.Y. Englewood Cliffs: Prentice Hall Inc.

CONTRIBUTIONS FOR FUTURE ISSUES

The next issue of Research Intelligence will contain articles on the study of history in education, approaches to curriculum development, some recent research in teacher education, a comment on the Bullock Report, a survey of educational research in Scotland and the continuation of Mick Youngman's article on statistical programmes.

Comments, opinions and articles for consideration for publication should be sent to the address given below. Articles should be typed on A4 single sheets with 1½" margin and with double spacing. Notes should be kept to an absolute minimum and placed at the end of the article. Tables and figures should be given on separate sheets and a clear indication of their placing given in the text. References should follow the style: AUTHOR (Date). TITLE. SOURCE (Book, Journal) and they should be given at the end of the article in alphabetical order. References within the article should be of the form (Author, date) The most usual length of article will be 2,500 words. Contributors wishing to submit extended papers are advised to consult the editor first.

Please address correspondence to:

Dr George Brown
(Editor: Research Intelligence)
University Teaching Services
University of Nottingham
University Park
Nottingham

SUBSCRIPTIONS

Research Intelligence is published by the British Educational Research Association and is available from Dr G Brown at the above address. Extra copies of this enlarged conference issue are available at the price of £2.00.

ACKNOWLEDGEMENT

The editor would like to thank the following for their help with the first number of Research Intelligence. Without it production would have been impossible.

Mrs. J. Clayton, Mrs. H. Hughes, Miss J. Clayton, Miss L. Clayton, Miss H. Hughes, Miss M. Gibson, Mrs. E. Mash, Miss S. Beck, Mrs. B. Keese.

FORTHCOMING CONFERENCES

Members are invited to forward details of any Conferences to the Editor, Dr. George Brown, University Teaching Services the University of Nottingham.

1. Association for the Study of Medical Education.

Annual meeting will be held at Nottingham University 22-23rd September, 1976. Further details from Conference Secretary, ASME Officer.
15Ob, Perth Road, Dundee.

2. British Psychological Society.

Annual Conference at University of York, 2nd-5th April, 1976.
Further particulars from the BPS 18-19, Albemarle Street.

3. "Future Psychologists, their task, training and needs." Conference at Holly Royde College, Manchester, 12th-15th February, 1976. Details from John Shaw, Staff Tutor, Extra-Mural Department, Manchester University.

4. International APLET Conference.

Association of Programmed Learning and Educational Technology are holding an International Conference on Individualised Learning at Dundee College of Education, 5th-8th April, 1976.
Further particulars from The Conference Secretary, ETIC Dundee College of Education.

5. Psychology of Language.

Conference at University of Stirling, 21st-25th June, 1976. Details from R.N. Campbell, Department of Psychology, University of Stirling.

6. BERA Annual Conference - Advance Notice. This will be held at Westfield College, London between the 7th and 10th of September, 1976. The Conference sub-committee is in process of drawing up the programme and would appreciate suggestions about topics or themes from members. Members who would like to give papers or organise a symposium should also contact the sub-committee, through the Chairman, Dr. M.P. Smith, Division of Humanities, Kingston Polytechnic.

BRITISH EDUCATIONAL RESEARCH ASSOCIATION

AIMS.

The broad aim of the Association is to encourage the pursuit of Educational Research and its application

TO: provide a critical professional framework for the discussion of problems and methods in educational research;

promote co-operation among research workers in various discpilines working in the field of education;

communicate educational research findings to all concerned with the study and practice of education

improve the training, facilities, and career structures of educational research personnel.

APPLICATION FOR MEMBERSHIP

Application for membership shall be made on forms provided by the membership secretary:

Dr. Derek Sharples,
Worcester College of Education,
Henwick Grove,
Worcester, WR2 6AJ.

Mercurius Bifrons ; Or , The Engliſh *JANUS* , Numb. 1.
The one ſide True and Serious, the other Jocular.

The True and Serious Intelligence :
Or, Occurences Domeſtick and Forraign.

From Thurſday *February* 24th to Thurſday *March* the 3d.

From *Kinſale.*

BY a Letter of the freſheſt Date, we have a Relation that there has lately been ſeveral conſiderable diſcoveries made

CONTENTS

[illegible] and a broad Sword, and then cutting [illegible] Bridles, they made with all ſpeed toward London.

His Majeſty being pleaſed to ſend for the Juſtices of the Peace of *Middleſex*, to acquaint them of his departure to meet his Parliament at *Oxford*, deſiring them to be Vigilent in his abſence, for that he had Intimation that ſome had armed themſelves above their Quality, and that they ſhould put the Laws in Execution againſt the Papiſts, for whatſoever ſtraits he ſhould be forced to, he would maintain the *Proteſtant* Religion; it being his ſole intent to govern by the Law as it is now eſtabliſhed, hoping his Subjects would yield obedience to the ſame, or words to this Effect.

A *Sales-man* and a *Taylor* playing together, at laſt fell to quarrel, whereupon the *Sales-man* ſtruck a Tobacco-pipe into his head 3 Inches deep, of which wound he dyed, and upon Inquiry, the Coroner and Jury on Munday laſt, found it Wilful Murther; whereupon it is reported the *Sales-man* is fled.

From *Chicheſter*. They write of the 28th. paſt, that his Grace the Duke of *Monmouth* being arrived there with a great Train of Gentry, was highly entertained by the Magiſtrates of that City, and that the Country Gentry from the adjacent Towns, have given his Grace a Viſit, who continues to recreate himſelf in, Hunting and the like Exerciſe.

From *Sterling*. We have Letters which ſay, His Royal Highneſs continues there, and that the *Mary* Yatch is every day expected, but they do not hear that his Highneſs intends to leave that Kingdom. Several of the Students are now at *Edenburg*, and it is confirmed from all hands, that in a ſhort time they will be recalled.

From *Brecknock*. We have a Relation that three Suns were ſeen there on the 23d. paſt, and that ſince the late *Comet*, there has appeared another more Prodigious then the former.

From *Dublin*. They write that a Veſſel coming with Paſſengers from thence, was lately caſt away upon that Coaſt of *Liver-Pool*, in which was Mr. *Barrer* and his Lady, with about 65 perſons more, who were moſt part drowned.

From *Plimouth*. On the 25th. paſt they write, That they have Intelligence from the Maſter of a *Dutch* Veſſel, driven in there by ſtreſs of weather, who lately came from *Guinny*, that Two *French* Veſſels were caſt away upon the *Guinian* ſhore, one of 400 Tuns, the other of 200 Tuns, having on Board them 1000 Blacks, and [illegible] weight of Duſt *Gold*, being bound for the *Weſt-Indies*.

From *Stafford*. We have an account that one *Evers* a *Popiſh* Prieſt, lately mentioned in His Majeſties Proclamation, was apprehended there, and committed to Priſon, where he now remains; and that it is greatly thought will make a conſiderable Diſcovery, having already declared many particulars that do much confirm and ſtrengthen the Evidence of Dr. *Oats*, and Captain *Bedloe*. Deceaſed

TUES-

A for the BERA Conference