AN ASSESSMENT OF MICROTEACHING IN THE CONTEXT OF THE GRADUATE TRAINING YEAR

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

TABLE OF CONTENTS

LIST OF TABLES

*Videotape copies of these skills Models may be obtained from the author.

APPENDICES

Appendix A Analysis of Teaching Skills

Appendix B Teaching Skills Appraisal Guide (first version)

Appendix C Teaching Skills Appraisal Guide (revised version)

Appendix D Appraisal of Teaching Skills (notes for lecturers)

Appendix E Teaching Skills Analysis sheet (question Technique)

Appendix F Four Basic Questioning Skills (notes for students)

Appendix G* Videotape Model script: Asking Questions

Appendix H* Videotape Model script: Dealing with Answers

Appendix I Student Evaluation form (Questioning Skills: Dealing with Answers)

Appendix J Student Questionnaire: term 2

Appendix K Final Student Questionnaire: term 3

Appendix L Lecturers' Questionnaire

Appendix M(a) Hints for Completing the Appraisal Guide (letter to tutors, October, 1972)

Appendix M(b) Supplementary note on Avoidance Behaviours (letter to tutors, December, 1972)

Appendix M(c) Reminder letter to Tutors, February 1973

Appendix N(a) Term 1 Briefing Letter to Students

Appendix N(b) Term 3 Briefing Letter to Students

Appendix O Ground Rules for Analysis and Coding of Student Lesson Excerpts

Appendix P Assessment of Recorded Lessons: Summer term, 1973 (rating form for tutors)

ACKNOWLEDGEMENTS

The author wishes to thank all who contributed to this research study, as students or teachers. In particular, he is grateful to the members of the Jordanhill College History department, whose zeal consistently outran their suspicion; to Mr Hugh Batten and Mr Roy Griffiths of the Department of Education at the University of Stirling, for their work on the coding instrument; and to his supervisor, Mr Donald McIntyre, magister exigens, collega gratus.

1. Introduction

Microteaching - a system of training in specific teaching skills through a series of scaled down teaching encounters - was introduced at Stanford University in the summer of 1963. Five years later, its use was reported (Johnson, 1968) at half of all the teacher training institutions in the U.S.A. Interest in the technique is now widespread in Europe, Africa and Australia. It has attracted a considerable number of research studies (reviewed by White, 1971; Griffiths, 1972, 1973; Brusling, 1974), aimed at assessing the effectiveness of the system as a whole and at analysing the inter-relationship of its component parts. Moreover, there have been numerous articles and reports describing experiments in microteaching at colleges of education and university education departments; experiments which, without having been evaluated in any rigorous or objective way, are nevertheless felt to be successful and appreciated by the staff and students involved. It is common at educational conferences to hear microteaching referred to - indeed almost taken for granted - as a promising innovation in teacher training. In Britain, at least, it appears to have achieved the same status as physical exercise. Most people approve of it, while not necessarily indulging in it.

Interest in microteaching developed at Jordanhill College of Education in the late sixties, following a presentation of Stanford University videotapes describing the system and a visit by a member of staff to the University of Illinois, where the technique was being extensively practised within the Teaching Techniques Laboratory. Initially, microteaching was associated in lecturers' minds with closed circuit television (which had been installed in the college in 1966) and its innovatory quality was thought of in technical rather than educational terms./

terms. Early experiments consisted in videotaping a short student lesson and subsequently reinforcing the traditional tutor critique by replaying portions of the tape. No systematic attempt was made to identify, analyse and concentrate upon specific teaching skills, nor was there much encouragement of student self-evaluation. Nevertheless, the interest aroused by these first sporadic trials, coupled with a growing awareness of the claims being made for microteaching as a means of bridging the gap between theory and practice in teacher training, prompted certain departments to move towards plans for incorporating an element of microteaching in their porfessional training curricula. In 1971, the History department expressed interest in a thorough-going experiment with their graduate students, to determine whether this technique could practicably and effectively be used to improve the quality of their training programme.

It could be argued that, by 1971, there was ample research evidence (reported, for example, by Allen and Ryan, 1969 and by Borg, 1970) indicating that teaching skills could be developed at least as effectively through microteaching as through traditional training techniques, with a considerable saving of student time; and that further research was justifiable only if directed to an examination of the most efficacious blend of elements and roles within this already proven system. It was considered at Jordanhill, however, that there were several strong arguments in favour of mounting a further basic research study:

(i) there were comparatively few research studies describing the use of microteaching with graduate students and no known research based on the graduate training year in Scottish colleges of education;

(ii) the practical problems of accommodating an element of microteaching/

microteaching within a short and intensive training period involving up to a hundred students needed examination;

(iii) comparatively little research had been done on the extent to which skills developed in microteaching could be retained and transferred to classroom teaching at a later date;

(iv) there was a need to study the results of microteaching in terms of "high inference" ratings of effective teaching and to assess the ability of student teachers to use skills appropriately, without unbalancing their teaching performance as a whole;

(v) it was important to examine the relevance of micro-teaching skills, identified in previous microteaching programmes, to a specialist subject area such as history and indeed to consider whether the experience of microteaching might lead to modification of accepted ideas about teaching methods in that subject.

In view of the rapidly increasing interest in microteaching, it was felt that evidence from the proposed Jordanhill research study would be valuable to other subject departments dealing with graduate students, in a variety of teacher training departments and institutions.

As a point of departure for planning, the research project accepted the five characteristics of microteaching listed by Allen and Ryan (1969):

- that microteaching is "real" teaching rather than role-playing, requiring the student to teach his own subject to real pupils;

- that/

- that the complexity of normal classroom teaching should be reduced by limiting the number of pupils and the length and scope of the lesson;

- that in each teaching encounter the student should concentrate on the practice of one or two specific skills, previously identified and analysed;

- that feedback on the microlessons should allow the student to evaluate his own performance in terms of the skills being practised;

- that in each sequence of microteaching the class size, environment, length of lesson, etc. should be controlled, thus freeing the student from the anxiety of facing unpredictable factors which might make it difficult for him to concentrate on the selected teaching skills.

It was also decided to adopt the structure of microteaching sessions developed at Stanford University, involving

- analysis and modelling of selected skills

- teaching of a short practice lesson

- critique based on feedback of performance

- replanning

- reteaching the same lesson to a different group of pupils

- critique based on feedback and discussion with supervisor.

The principal aim of the inquiry was formulated as follows:

to test the hypothesis that, within the graduate teacher training/

training year, history students given training through microteaching in specific skills will, in their subsequent classroom teaching, achieve performance in those skills significantly superior to the performance of a matched group of students who have been given no microteaching training.

In addition, the research study set out to examine:

(i) the feasibility of identifying and analysing a set of skills appropriate to history teaching in secondary schools and of creating an appraisal instrument, based on these skills, which could be used reliably by lecturers visiting students on school teaching practice;

(ii) the extent of agreement between assessments of students made by lecturers using an appraisal instrument on school visits and assessments of the same students based on a detailed quantitative analysis of recorded classroom lessons;

(iii) possible effects (beneficial or adverse) of microteaching training on students' classroom performance as a whole;

(iv) the practical problems involved, for students and lecturers, in organising microteaching sessions within the graduate training year;

(v) the reactions of students to the experiment;

(vi) the reactions of lecturers to the experiment.

The research was planned to extend over three sessions. In 1971-72, teaching skills would be identified and analysed; an appraisal instrument would be devised and tested for reliability; models of selected skills would/

would be prepared and recorded; and pilot microteaching sessions would be held, to explore the organisational problems involved. The main experiment would be run in 1972-73. Session 1973-74 would be spent in analysis and interpretation of the results.

The early policy and planning discussions with the History department brought to light problems which may well be characteristic of many attempts to introduce innovations into established courses within complex institutions. To begin with, there were organisational factors, beyond the department's control, which made it impossible to use microteaching as a form of initial training, on the pattern developed at the Universities of Stanford and Stirling. Graduate students at Jordanhill, divided into two main sections, do their teaching practice on a half term in/half term out basis throughout the session; so that, in the autumn term, half the student population, after a crowded week of introductory lectures, goes straight out to schools and remains on teaching practice for four days of each week until the mid-term change-over. Moreover, the students' time allocated to the History department during periods in college was already fully taken up with a carefully planned mixture of lectures (dealing with broad themes in history teaching), seminars (concerned with instruction in methods) and tutorials (allowing for discussion of practical problems arising in the course of teaching practice). Time spent on microteaching would mean cutting into what was felt to be a reasonably effective and well balanced course. Alternatively, if microteaching were to be inserted into periods of teaching practice, time would be taken from the activity commonly held by students to be the most valuable part of their training (Stones and Morris, 1972; Morrison and McIntyre, 1973). Not only that, but earlier experiments with microteaching, both at Jordanhill and at other Scottish colleges, had indicated that one tutor, working with one set of equipment, could guide no more than five or six students through a complete/

complete microteaching cycle in any one day. Thus, providing some fifty students with two cycles* of training in four basic skills would occupy the whole time of four tutors during a teaching practice period of five 4-day weeks.

Clearly, this would be an impracticable burden for a department with ten lecturers, unless some other major section of the course were to be curtailed. It was understandable that lecturers were reluctant to make sweeping changes of this kind in an established course without having first-hand evidence that these innovatory techniques were markedly superior to the accepted pattern of training. On the other hand, it was appreciated that such evidence would never be forthcoming if experiments were to be unduly restricted. In the end, a compromise was reached, offering two days of microteaching to a sample of ten students in each half of the second term. It was considered that microteaching, if it proved its worth, could feasibly be offered to approximately this number of students as a form of remedial training, and that this remedial function might well prove the most practicable way of integrating this technique within the curriculum.

Planning for effectiveness raised another problem with the History department. Assessment of students in teaching practice was on a 5-point scale, and based on the tutor's subjective impression of overall teaching ability. No formal attempt had been made to draw up a comprehensive list of teaching skills (though the Senior Lecturer had issued a handout to students, defining such basic skills as clarity and coherence of exposition, varying the stimulus and question techniques). Nor was any use made of detailed appraisal forms or observation schedules. The department claimed, however, that there was usually a fair measure of agreement/

---

*Research workers at Stanford University had claimed that two cycles of microteaching were the minimum for effective training in a skill (Allen and Cooper, 1968).

agreement when a student was rated by more than one tutor. The lecturers insisted that it would be naïve to assess "effectiveness" simply in terms of quantitative measures of technical skills; that "good teaching", being more than the sum of a battery of specified behaviours, involved judgement in using techniques appropriately and with discrimination; and that only a subject specialist could appreciate and assess the quality of a lesson in these terms. At the same time, they recognised that, to achieve reliability in their assessments, it would be necessary to move from purely subjective impressions towards a more rigorous and objective analysis of skills and towards some measure of agreement over performance norms.

It was acknowledged from the start that this research study should have an operational character, even though this might involve some reduction in precision and the acceptance of certain uncontrollable variables. The department wished to discover whether microteaching in the graduate training year could be effective at a level which could be accommodated without making major changes in established training procedures; and whether the extra effort and organisation entailed on lecturers and students could be kept within reasonable and practicable bounds. Finally, in making an assessment of microteaching, the college was concerned with performance in the classroom, not in the laboratory.

2. A Review of Relevant Research

At each stage of planning and developing the research study it was important to take account of previous findings. These can best be reviewed under headings which summarize the logical development of the project.

(a) Identification and analysis of teaching skills and behaviours

At the start of the experiment it was hoped to provide microteaching in a fairly wide range of teaching skills, in order to test the technique as fully as possible. It was therefore felt necessary to draw up a comprehensive list of skills applicable to history teaching at the secondary stage.

For the purposes of the study, it was found useful to adopt the distinction noted by Berliner (1969) and to group the extensive literature on teaching skills and behaviour under two heads:

(i) broad and generalised categories of behaviour, derived from studies of teacher characteristics, examination of teaching roles, analyses of classroom interaction and attempts to describe teaching in terms of strategies or sequences. Such studies have the merit of being comprehensive, but the categories are not always defined in the operational terms applicable to a teacher training programme;

(ii) lists of specific teaching skills, embodying measurable performance criteria. Such lists may have the advantage of precision, but some of them provide a sample rather than a complete account of teaching behaviour.

Under/

Under the first heading, for example, Ryans (1960) distinguished eighteen bipolar measures of teacher characteristics, such as responsive - aloof, stimulating - dull, systematic - disorganized. Clearly, this type of categorisation, however perceptive, could be ambiguous as a formula for practice and unreliable as a basis for assessment. More immediately relevant, perhaps, are the behaviour categories which have emerged from the numerous systems for analysing classroom interaction developed during the past twenty years. While these systems are essentially descriptive, research has indicated correlations between certain of the identified behaviours and pupil achievement, and this has prompted training institutions to use these systems as a basis for the development and practice of associated teaching skills. Thus Flanders (1963) identified seven categories of teacher talk, four of them "indirect" (accepting feeling; praising or encouraging; accepting pupil ideas; asking questions) and three of them "direct" (lecturing; giving directions; criticising or justifying authority). Meux and Smith (1964) analysed lessons into "units of discourse", which they categorised as defining, describing, designating, stating, evaluating, opining, classifying, comparing and contrasting, conditional inferring (from a given antecedent), explaining (sub-categorised under six headings) and directing and managing. Bellack and associates (1966), in their analysis of the linguistic behaviour of pupils in high school social studies classes, specified four types of "pedagogical move" - structuring, soliciting, responding and reacting - each with its "meaning", substantive or instructional, related either to the subject-matter being studied or to the social and managerial aspects of teacher-pupil interaction. Hough and Duncan (1970) listed nine "instructional behaviours" - initiation, clarification, response, solicitation, acceptance, confirmation/

confirmation, corrective feedback, positive and negative personal judgement - from which the teacher could select in developing the four basic teaching strategies of direct communication, teacher/ pupil interaction, independent pupil activity and group activity. De Landsheere (1970), basing his ideas on the work of Hughes (1959), analysed teacher behaviour into nine "functions", each of which was broken down into a number of operationally defined categories. The "functions of personal responses", for example, included

- welcomes a spontaneous participation
- invites pupil to tell or report about personal experiences out of school
- clarifies personal problem
- individualizes teaching.

Other headings noted by De Landsheere were controlling, impositive, content developing, positive and negative feedback to pupils, concretization and positive and negative affectivity. Some of the 46 functions, as defined, could reasonably be called "skills", and practised under microteaching conditions. De Landsheere's system was further developed by Bayer (1972), who distinguished two broad categories of communication functions (covering interaction between teacher and class) and instructional functions (in which the teacher is concerned with organizing work, imposing information, developing content and providing feedback).

Attempts at comprehensive analyses of teaching skills and behaviours show considerable variations in perspective and detail. One can choose, for example, the broad view of Millar (1972), who grouped teaching skills under the four general headings of motivating, explaining, soliciting and reacting; or that of Allen and Ryan (1969), who/

who distinguished five major categories of presentation, response, questioning, increasing pupil participation and creating pupil involvement. Closer analyses have been provided by Wood and Hedley (1968), who identified eight "discrete elements" of a lesson as introductions, expositions, demonstrations, discussions, questions, recapitulations, conclusions and assignments; and by Briggs (1968), who defined basic instructional functions as

- formulating goals for instruction-learning units in terms of the desired or required performances
- gaining and directing attention of learners
- presenting and displaying instructional content or stimulus materials
- eliciting or stimulating learning responses
- providing feedback and reinforcement
- orienting and regulating successive efforts as learning occurs and becomes more complicated
- assessing and assisting the student in evaluation of his own performances and otherwise confirming the learning responses.

More detailed still is the list of teaching activities draen up by Komisar (1966), including introducing, demonstrating, citing, reporting, conjecturing, confirming, contrasting, explaining, proving, justifying, explicating, defining, rating, appraising, amplifying, vindicating, interpreting, questioning, elaborating, identifying, designating and comparing.

A study of lists such as these reveals differences not only in perspective/

perspective and detail but in ways of looking at the act of teaching. Different definitions of skills and behaviours appear to overlap. A behaviour defined with relative precision, such as Briggs' "presenting and displaying instructional content or stimulus materials", would seem to fit equally well into either "motivating" or "explaining", as defined by Millar. Moreover, Cooper (1968) contends that "skills will differ according to subject field, grade level and a host of other variables" and goes on to assert that "there is no one set of technical skills which is better than another set". Faced with this rather daunting freedom of choice, anyone embarking on a programme of microteaching understandably turns for guidance to lists of skills already used as the basis for microteaching exercises in other institutions. At Jordanhill, a number of such lists was examined and information was drawn from three sources in particular: the Stanford University programmes, the Minicourses developed by the Far West Laboratory for Educational Research and Development, and the work done at the University of Sydney.

At Stanford University, one of the first skills to be identified and practised in microteaching sessions was set induction, involving methods of introducing the subject matter of a lesson in a way calculated to capture attention, stimulate interest and provide cues for the understanding of what was to follow. As the work at Stanford developed, the skills were defined more precisely and each skill was broken down into a number of teacher behaviours. It was recognized that certain skills such as varying the stimulus and lecturing (referred to as integrative skills) involved combinations of other skills. This increase in precision was not associated with any rigid dogmatism. Skills were re-formulated from

one/

one year to the next, through consultation between students and supervisors, and the approach remained exploratory. Allen and Ryan (1969) listed the following fourteen skills developed at Stanford University; skills "that can be applied at many levels, for teaching many different subjects":

- stimulus variation
- set induction
- closure
- silence and non-verbal cues
- reinforcement of student (i.e. pupil) participation
- fluency in asking questions
- probing questions
- higher order questions
- divergent questions
- recognizing attending behaviour
- illustrating and use of examples
- lecturing
- planned repetition
- completeness of communication

Ward (1970), reporting on microteaching programmes in 141 American teacher training institutions, stated that the technical skills generally considered to be most important were, in order of priority, probing, reinforcing, questioning, higher order questioning and establishing/

establishing set; while the skills most frequently practised were questioning, establishing set, reinforcing, use of examples and varying the stimulus. Less than a third of the institutions, however, were using written definitions or perceptual models of specific skills; a situation which perhaps reflected lack of experience rather than lack of conviction.

In California, at the Far West Laboratory for Educational Research and Development, self-instructional microteaching programmes, known as Minicourses, were developed from 1967 onwards for teachers in service. The selection of specific teaching skills was influenced by the work already done at Stanford University, but the final definition, analysis and grouping of skills resulted from a great deal of research and field testing. Eighteen Minicourses have been planned and are at various stages of development and distribution. The teaching skills cover questioning at different grade levels, developing oral language, teaching reading, individualising mathematics instruction, organizing independent learning, role-playing, discussion techniques and heuristic approaches. Within each course, training is given in a number of specific skills. Thus, "Individualising instruction in Mathematics" is broken down into

- using verbal praise to reward correct responses
- asking prompting questions to increase pupils' active involvement in the tutoring process
- asking general diagnostic questions
- questioning techniques appropriate to number operations and verbal problems
- estimating/

- estimating an answer before using a computational algorithm

- estimating techniques appropriate to number operations and verbal problems

- assigning examples (for evaluation and practice)

- having pupils correct their own work

- having pupils tutor each other

- assigning monitors for management tasks

This breakdown of skills illustrates a number of important concepts developed by Borg and his associates at the Far West Laboratory, namely

- basic skills such as reinforcing, prompting and questioning are re-defined in relation to subject area (mathematics) and grade level (elementary)

- the skills are constructed as a sequence; for example, the techniques that help to develop pupils' understanding of mathematical procedures would logically be used after the skills of diagnosing grasp of concepts

- management functions (assigning monitors, etc.) are included along with didactic skills

- an attempt has been made to define skills appropriate to group and individual work as well as to class teaching.

In/

In Australia, at the University of Sydney, Turney and associates developed microteaching projects from 1969 onwards with groups of both primary and secondary education students. Once again, the skills formulated at Stanford University provided the point of departure; but fresh analyses were made and a number of new skills were added, particularly at the primary level, for example

- encouraging creativity
- guiding discovery
- individualising instruction
- developing concepts.

At the secondary level, Owens and Hatton set skills training in the context of defined lesson tasks, moving (between Term 1 and Term 3 sessions) from basic to integrative skills and from didactic to discursive strategies. The experimental work of these early sessions is now being consolidated in the "Teaching Skills Development Project" (Turney, Clift, Dunkin and Traill, 1973). 81 teaching skills have been identified and classified under the headings of

- motivational (7 skills)
- presentation and communication (29 skills)
- questioning (9 skills)
- small group and individual instruction (8 skills)
- developing pupil thinking (10 skills)
- evaluative (5 skills)
- classroom/

- classroom management and discipline (13 skills)

It is recognized that the skills are not necessarily discrete, nor are the broad categories mutually exclusive. The authors of the classification system maintain that "since the skills are all part of the complex teaching act, their frequent overlapping and interlocking are to be expected".

One aspect of behaviour which it was hoped to influence through microteaching was the percentage of teacher talk, particularly in lesson sequences where the teacher was attempting to elicit ideas and encourage discussion. Studies over the past sixty years (for example, those reported in Borg et al, 1970) confirm consistently that teachers tend to talk far more than pupils. Stevens (1912) analysed class discussions led by teachers in New York high schools and found that the teachers talked for approximately two thirds of the time. Corey (1940), analysing talk in six high school classrooms during a period of a week, found that the teachers provided 64% of the word total. A word count of elementary school lessons by Floyd (1960) produced a 71% figure for teacher talk. Flanders (1963), summarizing earlier research in interaction analysis, formulated his "rule of two thirds" - two thirds of observed lessons were taken up with talk, two thirds of this talk was by teachers and two thirds of the teacher talk consisted of lecturing. Bellack (1966) transcribed the content of sixty social studies lessons in fifteen classes and found that 72.1% of the lines of transcript represented talk by the teachers. Bayer (1972) counted the number of sentences spoken in extracts from 45 lessons given by 15 elementary school teachers in arithmetic, reading, science and social studies. He found that 67.5% of the sentences were spoken by the teachers, either to the class as a whole or to individual/

individual pupils.

Griffiths (1973) has argued that "the effectiveness of microteaching may be increased by attending to the development of component skills of teaching which are defined in such a way as to relate more closely to appropriate educational objectives within a particular area of educational content" and has suggested that "for example, the ways in which higher order questions are used in the teaching of maths and history may differ so widely as to necessitate distinctive formulations of the relevant skilled behaviours". No record has been found of any previous attempt to identify technical skills specifically related to history teaching, though Coltham and Fines (1971) have developed a taxonomy of educational objectives for the study of history. The framework of the taxonomy is constructed under the headings of

- attitudes towards the study of history
- nature of the discipline
- skills and abilities
- educational outcomes of study

The skills and abilities are those which the pupils should acquire and are listed as vocabulary acquisition, reference skills, memorisation, comprehension, translation, analysis, extrapolation (a more creative mental process than extrapolation as defined by Bloom (1956) in his taxonomy of educational objectives), synthesis, judgement and evaluation, and communication skills. While the skills are defined in terms of pupil behaviour, the list is of great use to the trainer of history teachers, since the behaviours imply headings under which teaching skills specific to the study of history might/

might be developed. For example, the pupil skills of "formulating what is interesting, puzzling, etc. about a piece of evidence or secondary source material" and of "identifying inconsistencies and bias" in such material suggest the possibility of identifying and defining corresponding teaching skills specifically related to work with historical documents.

The Jordanhill History department attached great importance to training in questioning skills, and the project eventually concentrated on questioning to the exclusion of other skills. Concern with the purpose and the cognitive level of student questions led to an examination of previous attempts to establish question categories. In recent years, many such categories appear to be based on the work of Bloom et al (1956) and Guilford (1956). Bloom established a taxonomy of educational objectives which related questions to the type of cognitive process required in the response, moving upowards through the levels of knowledge, comprehension, application, analysis, synthesis and evaluation. Guilford distinguished five categories of questioning related to five classes of thinking: memory, cognition, convergent thinking, divergent thinking and evaluative thinking. These categories were reformulated by Gallacher and Aschner (1963), who listed four types of questions

- cognitive memory (asking for recall, with no additional thinking)

- convergent questions (requiring analysis of data)

- divergent questions (calling for implication, prediction, imagination)

- evaluative questions (involving structured judgements)/

judgements).

Taba and Hill (1965), working with elementary school pupils in the area of social studies, related questioning to different levels of mental operation grouped under three broad headings of concept formation, interpretation of data and application of principles. Bellack (1966) classified discussion lesson questions as those requiring fact-stating, explaining, analysis or evaluation. Shipley et al (1968) distinguished three broad categories of factual questions, problem questions (requiring reasoning) and opinion (or open) questions. Barnes (1969) used similar categories in analysing teacher questions, but made a distinction between "closed" and "open" reasoning, and between open questions which involved reasoning and those which called rather for imaginative, creative thinking. He also added a category of "social" or procedural questions. Tinsley and Davis (1969), in their work with social studies student teachers, retained the Gallacher and Aschner categories, but added logical thinking and reflective thinking.

Hough and Duncan (1970) found it useful to consider questioning in terms of the teacher's purposes in developing reciprocal communication between himself and his pupils. Questions could be used for the purposes of diagnosis, review and measurement of learning (to test the extent of the pupils' knowledge and understanding); reinforcement (where the teacher is more concerned with rewarding correct responses than with diagnosing weaknesses and errors); and stimulation of either closed or open thought processes.

The Far West Laboratory team (Borg, Gall et al, 1970, 1971) based their Minicourse questioning categories on Bloom's taxonomy. Batten (1972), in a development of Minicourse 9 on Higher Cognitive Questioning/

Questioning, defined a category of "lower order synthesis" questions, to classify those teacher questions which "intend the pupil to venture an opinion not based on criteria other than personal preference".

The classification systems listed above are sufficiently broad to be related to any area or level of the curriculum. Gall (1972) advocated more detailed classifications applicable to specific subject specialisms and suggested that identification of distinctive question types relevant, for example, to mathematics tutoring, scientific concepts, role-playing in social studies and discussion of controversial issues might provide student teachers with more precise guide-lines for improving their questioning techniques than broad categories based on cognitive processes. Gall also drew attention to the need for other kinds of question categories not covered by cognitive taxonomies; for example, questions to stimulate curiosity and a sense of inquiry, and follow-up questions after initial responses. Different types of follow-up questions (prompting, probing, re-direction) have, in fact, been defined in the Far West Laboratory Minicourses and in microteaching courses at the University of Stirling and elsewhere.

The need to vary the level of questioning and to increase the proportion of "higher order" questions (covering Bloom's categories of analysis, synthesis and evaluation) found support in research literature. Stevens (1912) reported that in high school classes two thirds of the teachers' questions were limited to recall of facts. Haynes (1935) found that 77% of the questions asked in history lessons with 10-11 year old pupils were to test recall. More recently, Floyd (1960) at the primary level and Gallacher (1965) and Davis and Tinsley (1967) at the secondary level classified/

classified approximately half of the teachers' questions as factual. Ninane (1969) and Bayer (1972), using Bloom's taxonomy to analyse primary school lessons in a variety of subjects, found that higher order questions were limited to 9% (Ninane) and 14% (Bayer).

(b) The relation of teaching skills to "effectiveness" in terms of learning behaviour.

By comparison with normal teaching practice, microteaching involves spending a good deal of student time on detailed aspects of the teaching process. The Jordanhill History department was therefore concerned that the skills chosen for the study should be associated with "good" teaching; and, at a later stage of the project, when the decision was made to concentrate on question techniques, it was felt important to select those questioning skills which research had shown to be related to learner achievement. The subtlety and complexity of the links between teacher behaviour and pupil gains have been commented on by Gage (1968), Rosenshine and Furst (1971) and Morrison and McIntyre (1973). Skills which are effective with one age group may not be appropriate to another. Pupils of varying intelligence and personality may react in different ways to the exercise of a particular skill. A tactic such as open questioning may produce rich rewards in a lesson on English literature, but lead to vagueness and confusion in modern language teaching. Rosenshine and Furst maintained that "the relationship between the teacher behaviours advocated by educational experts and the consequent learning by students has not been thoroughly investigated" and pointed out that this lack of evidence could lead different training institutions to advocate opposing performance/

performance criteria; for example, the Far West Laboratory's Minicourses trained teachers to reduce their repetition of pupil responses, while the North West Regional Laboratory programme, based on Flanders' interaction analysis categories, encouraged more repetition of pupil responses as a component of the preferred "indirect" style of teaching.

The research which has been carried out on the relationship of teaching skills to pupil achievement (and Rosenshine and Furst (1971) reviewed some fifty studies in this field, dating back over twenty years) has not always obtained significant results, and the reviewer has to interpret a patchwork of evidence in which not all the pieces match. Research carried out by Gage (1968) indicated that the factors most usually associated with effective exposition (as measured by pupils' performance on tests) were the use of "explaining links" (underlining the how and why or the effect of something) and the "rule-example-rule" pattern of exposition. Unruh (1968) found that teachers rated high on planning and reorganization, knowledge of material and presentation skills tended to produce more learning in their pupils. Rosenshine (1970) summarized six studies employing high inference measures of teaching (i.e. measures "which require considerable inferring from what is heard or seen in the classroom to the labelling of the behaviour", as distinct from low inference measures "which require the observer to classify teaching behaviours according to relatively objective categories"). There was consistent evidence that teacher enthusiasm correlated positively with pupil achievement. Associated low inference variables, identified in some of the studies, included frequency of questioning (requesting interpretations, opinions and facts), praise and encouragement, and variation in gesture, movement and voice. Flanders (1970) reported/

reported research findings indicating that pupils learned better from "indirect" teaching (categories 1-4 of his interaction analysis system) and concluded that "when pupils have opportunities to express their ideas, and when these ideas are incorporated into the learning activities, then the pupils seem to learn more and to develop more positive attitudes towards the teacher and the learning activities." The FIAC categories 1-4 can be associated with the Stanford skills of reinforcement of pupil participation, fluency in asking questions and probing. However, other research studies investigating the relationship of these skills to pupil achievement have been largely inconclusive. Turney et al (1973) note that, among fourteen recent studies of teacher reinforcement, ten found no significant relationship between praise and achievement, one found a negative relationship and only one reported an unqualified positive relationship.

Fluency and frequency of questioning are broad-band behaviours which require further analysis. Hunkins (1967, 1968) studied the relationship of question type (defined in terms of Bloom's taxonomy of educational objectives) to pupil achievement, working with social studies classes of 10-11 year old children. He found that pupils who had been questioned predominantly at the level of analysis and evaluation scored significantly higher on a post-training test (which presented multiple choice questions designed to cover the whole range of the taxonomy) than pupils who had been subjected to questioning predominantly at the knowledge level. In detail, the analysis-evaluation group were significantly superior in the application and evaluation sections of the test and equal to the knowledge group in all other sections. These findings would appear to conflict with the evidence of studies by Wright and Nuthall (1970) and Francis (1971), reported by Nuthall and Church (1973)/

(1973). In these studies, pupils taught by an approach emphasizing open-ended questioning and stimulus of "ability to think" responded more ambitiously but learned a smaller amount of factual content than pupils taught by "achievement-oriented" methods with an emphasis on direct learning.*

Other studies reviewed by Nuthall and Church suggested that

- variations in the pattern of addressing questions to pupils had no effect on learning

- positive teacher reaction to pupil responses produced a significant improvement in learning

- increasing the amount of time spent on given content produced greater changes in achievement than variations in teacher questioning and reacting behaviours when coverage of content was held constant

- pupils questioned frequently on given content learned more than pupils who merely listened to responses from others; but these listeners in turn learned more than pupils in classes where the content was provided as straight exposition, without questioning

- telling/

---

*Professor Flanders, in a lecture at the University of Stirling (1972), stressed the danger of over-generalization in discussing "pupil achievement". Research evidence indicated that achievement in terms of memory was associated with a direct style of teaching, whereas measures of pupil creativity correlated positively with indirect teaching.

- telling pupils the answers to questions which they could not immediately answer produced less learning than persistence with prompting and redirection of questions.

In their summary of correlational studies, Rosenshine and Furst (1971) found some evidence of association between pupil achievement and

- frequency of questioning
- structuring of questions (i.e. establishing content and providing guide-lines)
- varying the types of questions, in phrasing and in cognitive level
- the use of probing questions, as a follow-up to initial responses.

There would appear to be no direct, linear relationship between pupil achievement and the frequency of questions at a specific cognitive level. The writers emphasize that frequency of use may not be a sufficient measure in relation to skills such as questioning. Pattern, sequence and context may be equally significant. They stress, too, the importance of keeping pupil differences in mind and the danger of generalising from statistics of mean performance.

(c) Effectiveness of microteaching in comparison with other training methods.

The initial reaction of any teacher training specialist involved in/

in the introduction of microteaching can be summed up by two blunt questions: "Does the technique work?" and "Does it work better than other training methods?" There is a good deal of evidence that microteaching is effective in the sense that the trainees' performance on a selected teaching skill improves significantly between the initial teach and final re-teach sessions. In the first few years of microteaching at Stanford University, performance was assessed by use of the Stanford Teacher Competence Appraisal Guide (designed to appraise a broad range of teaching behaviours), applied to short "laboratory" lessons. Bush (1966), reporting on the first Stanford microteaching clinic, noted significant changes in the three skills practised by the experimental group, with performance at a higher level than that of the conventionally prepared control group. In their report of the 1965 Stanford clinic, Fortune, Cooper and Allen (1967) claimed that microteaching had produced highly significant changes on nine items of the S.T.C.A.G., even though only five skills (initiating, presenting, consolidation, monitoring and evaluation) had been practised that year. In another report, Allen and Fortune (1966) stated that a microteaching group with ten hours' training were rated significantly higher (using the S.T.C.A.G.) on teaching effectiveness than a group which had received 25 hours of conventional training. This study was replicated by Kallenbach and Gall (1969) with a group of 37 students at San Jose State College. During a summer clinic, a control group were given 50 hours of conventional teaching practice, while the experimental group carried out seven hours of microteaching. Allowing for travelling and preparation time, Kallenbach and Gall claim that the microteaching occupied the students for one fifth of the time spent by the teaching practice group. Performance of both groups was assessed

- through/

- through a 5-minute microteaching lesson at the end of the training period

- through a ½-hour observation of classroom teaching in the following autumn (scored by the S.T.C.A.G. and an Instrument for Observation of Teaching Activities)

- through a 1½ hours' observation of classroom teaching in the following spring (scored by I.O.T.A.).

No significant differences between groups emerged from any of these assessments, a finding that led Kallenbach and Gall to claim superiority for microteaching on the grounds of economy of time. However, the study raises a number of queries, for example

- no comparison is made of the time spent by tutors involved with the two groups

- the estimate of 75 hours of travel and preparation time for the control group, as against 14 hours' preparation time for the microteaching group, may reflect conditions at San Jose, but can scarcely be accepted as the norm for other institutions

- no mention is made of the number of skills practised in the seven hours of microteaching, but it is difficult to imagine that the training could match the range of experience to be obtained from 50 hours of teaching practice

- one must question the validity and reliability of/

of S.T.C.A.G. assessments of overall teaching competence, based on a 5-minute lesson taught under laboratory conditions

- the detailed list of mean scores raises a number of unanswered questions. The experimental group, for example, improved slightly through the training period, but showed a more marked improvement in the autumn ratings. The control group actually dropped back in performance during the teaching practice period, but achieved a sizeable advance in the autumn.

Several other studies have produced evidence of positive results. Bell (1968), working with home economics teacher trainees, found that a group which had received microteaching tuition in addition to normal teaching practice showed significant gains in performance between diagnostic and criterion lessons and also performed significantly better than a control group which had undergone normal teaching practice alone. Berliner (1969) gave microteaching training in higher order questioning to 120 graduates, using different modelling procedures. All groups showed significant gains in the number of higher order questions which they employed. Borg (1970) reported the results of field testing Minicourse 1 (Effective Questioning - Elementary Level), developed by the Far West Laboratory. Analysis of classroom lessons taught immediately before and after training showed significant gains in ten of the eleven skills covered by the course. These gains were maintained in lessons taught and analysed four months after the course, with significant loss in only one skill (prompting) and significant further gains in the skills of clarification and avoiding the repetition/

repetition of questions. Turney (1970), in an experimental four-week microteaching course with students at the University of Sydney, noted "quite substantial" improvement by a number of students as measured by a comprehensive scale of teaching competence applied at the beginning and end of the course.

All this evidence stems from projects in which the cycle of work was built around the act of practising a skill with a small group of pupils or peers. Two studies investigated the effectiveness of microteaching when the act of teaching was omitted from the cycle. Britton and Leith (1971) gave training in set induction to 28 student teachers. 15 of the students had an opportunity to practise the skill, while the remainder merely watched the teaching and participated as "pupils". Students who had practised the skill were rated higher (though not significantly so) than the students who had merely observed. The combined average ratings of these two groups was significantly higher than that of a control group which had received normal teaching practice while the microteaching course was in progress. These findings are challenged to some extent by the work of Kissock (1971). He gave eight hours of instruction on higher order questioning to 69 students, using video and written models of the skill. Half the group practised the skill in four 5-minute micro lessons, while the remainder did no teaching. At the end of the instruction period, a post-test showed that both groups had improved significantly since the post-test, but that the teaching group used significantly more higher-order questions than the observation group. However, a retention test four weeks later showed no significant difference between groups; the teaching group having regressed while the control group maintained their performance. Kissock suggests that the teaching act within the cycle "may work primarily/

primarily as a means by which a person can demonstrate what he has learned from the instruction program, but not as a vehicle for learning the skill itself".

To be acceptable to a trainer of teachers, a training technique must prove its worth within the constraints under which he is obliged to operate. The Jordanhill History department was therefore particularly interested in evidence that microteaching was effective in the context of the graduate training year. Few studies within this kind of framework came to light. Owens and Hatton (1970) worked with graduate students on a Diploma of Education one-year course at the University of Sydney and integrated microteaching into a problem-centred, "situational teaching" programme. Two sessions of microteaching were run: the first in Term 1, introducing basic skills such as reinforcement and questioning, and the second in Term 3, concentrating on integrated techniques such as encouraging inferential and divergent thinking. School assessments of student teaching indicated gains in the use of some basic skills, and student reaction was favourable. Gregory (1971) introduced microteaching into the first term of the graduate Certificate of Education course at the University College of Rhodesia. Four skills were studied, with each of the 71 students practising each of the skills twice. No objective assessment was made, but a questionnaire showed that three-quarters of the students felt that the course had helped them substantially in their teaching. Heaps (1973) carried out a similar experiment with graduate students taking a Diploma in Education course at the University of Nairobi. Ten lectures were given on basic skills and there were ten practical sessions. Again, no objective assessment was made, but it was generally felt that the exercise was very successful.

Microteaching is still an experimental and exploratory technique and it is therefore dangerous to generalize about its effectiveness in relation to other training methods. For this reason, much recent research (some of it reviewed below) has concentrated on an examination of the relative effectiveness of different elements in the process and on the most effective blend of components within the cycle. It is important, too, to take into account the work of researchers such as Koran J J (1968, 1971) and Koran M L (1969, 1971, 1972) on appropriate instructional methods to fit varying trainee characteristics. Their findings indicate, for example, that more able students learned better from written models of skills, while less able students achieved more from video models. Again, students of varying ability appeared to react differently to varying types of video model presentation, with students of relatively low ability benefiting most from models concentrating attention on both teacher and pupils, and limiting examples to positive use of skills, whereas negative examples ("how not to do it") proved effective training devices with students of higher ability.

(d) Transfer of microteaching performance to classroom practice.

Both students and training staff tend to be sceptical about the likelihood of transferring skills acquired in the "artificial" setting of a microteaching encounter to the more exacting environment of the classroom; and it is possible, as Berliner (1969) has stated, that "through concern for reducing the complexity of the classroom and the length of a lesson...a situation yielding little transfer effect to the classroom may have been produced"./

produced". Berliner goes on to suggest that "a concentrated effort should be made to determine the magnitude of transfer from artificial micro environments to real macro environments...Without this kind of information, it is not known if training teachers in specific teaching skills is an academic exercise or a program having genuine impact on education". Unfortunately, the evidence from research in regard to transfer is both sparse and inconclusive. Kallenbach and Gall (1969) carried out classroom assessments of overall teaching competence in the autumn and spring following a summer microteaching clinic, but no clear picture emerges from their findings. At the start of the clinic, the control group were significantly superior to the experimental group on a diagnostic test. Thereafter, no significant experimental/control differences were found. The mean performance of both groups appeared to rise in the autumn assessments and fall back again in the spring to the level of the summer criterion test performance. The Far West Laboratory team (Borg et al, 1970) made a more thorough attempt to measure transfer. In their field testing of Minicourse 1, for example, performance on the component skills was measured from classroom lessons taught immediately after the course and subsequently after a 4-month interval. Significant gains in performance noted in the post-course test lesson were sustained in the delayed assessment for all but one of the skills. It should be noted, however, that the post-course assessments were of lessons strictly controlled in terms of subject matter and selection of pupils.

There would appear to be no other findings comparable in scope with those of the Far West Laboratory. Assessments in the Britton and Leith (1971) study were based on classroom teaching a term after the microteaching course, but it is difficult to judge the significance/

significance of any relationship between practice of a single skill (set induction) and evaluation of general teaching competence. Wright (1973) provided training in set induction, stimulus variation, and closure for a group of 23 third year students at Hamilton College of Education, with initial and final assessments made in the classroom by a team of four independent tutors, on the "traditional basis of general impression". Mean performance overall and for each of the three skills rose significantly, approximately three-quarters of the students showing improvement. However, no control group was involved in the study. Kelly (1973) gave microteaching training to small groups of Coventry College of Education students, to increase their use of "reasoning" questions, and based pre- and post-assessments on classroom lesosns. Initial performance showed no significant differences between control and experimental groups, but the experimental groups improved significantly in post-test performance, as measured by the number of reasoning questions asked, while the control group remained static. All training, including the initial and final tests, was compressed into a three-week period.

Kelly's findings appear to conflict with those of two American studies which also used low inference measures to assess transfer effects. Brashear and Davis (1970) found significant difference between microteaching and control group in only one out of eighteen OScAR 5v categories which were used to measure classroom performance subsequent to a microteaching training programme. Copeland and Doyle (1973) trained a group of social studies students in three questioning skills over a six-week period and, seven weeks later, measured classroom performance of the trainees and a control group, using a rate per minute coding system. No

significant/

significant differences were found between control and experimental groups, either in the rate of total questioning moves or in the rate of moves in the specific questioning skills. The authors conclude that, "although laboratory skill training would appear to facilitate skill learning among teacher trainees, such training alone may not be a sufficient condition for effecting skill performance in the classroom" and suggest further research into aspects of student teaching experience which may support or inhibit the effects of laboratory training.

(e) The practice of microteaching.

The structure of the training programme at Jordanhill was determined partly by practical constraints and partly by the experience of other projects. Guidance from research studies was sought under the following heads:

(i) Modelling

In the context of microteaching, a model may be defined as a demonstration of a teaching skill or behaviour. A model may be positive ("how to do it") or negative ("how not to do it") and it may be presented in film, videotape, audiotape or tape/slide format (referred to hereafter as perceptual modelling) or a written lesson transcript or a set of written instructions (referred to as symbolic modelling). A good deal of research (usefully reviewed by Turney, 1973 and Brusling, 1974) has been

carried/

carried out on modelling procedures, much of it seeking answers to three basic questions:

- does the provision of models significantly enhance the effectiveness of the microteaching cycle?

- are perceptual models superior to symbolic?

- what is the importance of cueing (i.e. comments on the occurrence and use of skills, superimposed on the picture, incorporated in the sound track, supplied by a supervisor or added to the transcript), to provide discriminatory training for the students?

Evidence on the first of these points can be gained from several studies. Salomon and McDonald (1969) found that teachers who viewed their own teaching performance on videotape, without the provision of models or supervisory criticism, reacted according to the degree of their expressed satisfaction in their performance. The authors concluded that self-viewing on videotape will not lead to any desirable attitudinal or behavioural changes unless the viewer has previously been given, and has accepted, a desired standard or model of performance. Koran M L, McDonald and Snow (1969) used modelling as an independent variable in training students to ask analytic questions. Students exposed to the models proved significantly superior in their use of the skills to a no modelling control group. Lange (1971) prepared a model of indirect teaching relating to the discussion of a story and showed it to a group/

group of twenty elementary student teachers, while a comparable group viewed a film bearing no relationship to the target skill. In a test lesson two days after viewing, the experimental group showed a significant change towards indirectness, whereas the control group showed no change. Bierschenk (1972), in a study similar in some ways to that of Salomon and McDonald, found that, in the absence of modelling, microteaching experience appeared to have no demonstrable effect on the way in which student teachers perceived and evaluated their own teaching or the behaviour of their pupils. In a micro-teaching training programme for teachers of pre-school children, Rutherford (1973) found that the incorporation of modelling, with or without video feedback on performance, into the training cycle produced improvements in the use of the target skill; whereas a cycle which included video feedback but excluded modelling produced no such improvements.

A comparison of perceptual and symbolic modelling was made by Orme (1966) at Stanford University, in a programme designed to train students in the use of probing questions. The modelling variables were

- studying written instructions on the use of the skill (symbolic)

- viewing a videotape model on one's own (perceptual)

- viewing a videotape model with a supervisor present to provide cues (perceptual)

Orme/

Orme found that the perceptual modelling treatments led to significantly greater gains in probing techniques than did symbolic modelling. Supervisor cueing also influenced performance and Orme concluded that "the most effective variable for describing a desired behaviour appears to be a modelling condition in which the behaviour is portrayed and in which the subject views the model's performance while being cued by an experimenter on the significant aspects of the model's behaviour". This conclusion was supported by a study carried out by Young D B (1967), comparing various types of perceptual models with symbolic modelling. Again, the perceptual modelling technique proved superior. The study by Koran M L et al (1969), referred to above, also found that models presented on film were more effective than symbolic models in the form of a verbatim transcript of the film sound track. This finding is of interest particularly since the dependent variable was a verbal questioning skill.

Several studies, however, have detected no significant difference between the effects of perceptual and symbolic models. This finding emerged from two evaluations of training in higher order questioning, carried out by Allen, Berliner, McDonald and Sobal (1967) and by Gall (1972). Phillips (1973) reached the same conclusion on the basis of a project to train social studies students in classifying questions. Both experimental and control groups were given handouts describing the classification system. The experimental group then viewed a videotape showing a teacher asking questions at different levels, while the control group were given a transcript of the videotape/

videotape sound track. Another study on similar lines was carried out by Jacobsen and Gerlach (1973) in the context of a training programme on questioning skills. No significant differences between the perceptual and symbolic model groups were found, either on a micro lesson post-test or in follow-up observations conducted during subsequent teaching practice.

The studies already cited, by Orme (1966) and Young (1967), support the value of cueing in the modelling procedure. Claus (1969) also investigated the function of cueing in relation to both modelling and feedback components of a microteaching cycle designed to train elementary school teachers in higher order questioning skills. She found that cued modelling was significantly more effective than non-cued modelling in producing the desired behaviour change, but that cueing appeared to have no effect on the value of self-observation of videotaped lessons. A detailed analysis of the performance figures raises certain queries in relation to these findings - for example, the only group to show consistent progress in higher order questioning throughout the four lessons of the training cycle was the group exposed to non-cued modelling and cued feedback - but the largest pre-test to post-test gain in the number of higher order questions asked was achieved by the group which received cues in modelling but no cues in feedback.

(ii) Pupils v. peers

The/

The use of students to play the role of pupils in micro-teaching sessions effects obvious savings of time and expense. Evidence from a number of institutions confirms that students prefer the "reality" of working with pupils, but whether this preference helps them to acquire the skills more effectively is uncertain. Wood and Hedley (1968) write of a "diminishing return" when peers were used in place of pupils in microteaching with Canadian students. An interim report by Levis (1973) on work with pupil and peer groups at Macquarie University suggests that, even where no significant difference could be found in performance, students on the whole preferred to teach pupils, though they recognized that peer groups could provide more effective feedback.

(iii) The microteaching cycle

A number of questions arise under this head, concerning the optimum length of each teaching encounter, the appropriate number of pupils or peers, the most effective interval between teaching and feedback, and between teach and re-teach, and the necessity of retaining the re-teach as an element in the cycle. Decisions on these matters are often prompted by the exigencies of student numbers and time-tables, but a certain amount of evidence is available from research studies.

The approach to the mechanics of the microteaching process was kept flexible by the development team at Stanford University/

University. Allen and Clark (1967), for example, maintained that "a precise definition of the microteaching context can vary according to the purposes and resources of the user. Some of the variables which can be adjusted include lesson length, number of pupils, types of pupils, number of re-teaches, the amount and kind of supervision and the use of videotape". However, a 5-minute period was established as the standard length for individual micro lessons and 20 minutes for lessons taught as part of a unit by a team of students. There would still appear to be no definitive research evidence concerning optimum lesson length or number of pupils, though several studies (Gregory, 1971; Heaps, 1973; Wright, 1973) report that students felt that a 5-minute lesson was too short to be useful. However, a considerable majority of students in a project carried out by Brusling (1974) thought that 5 minutes was "just right".

A comparison of different delays in feedback, in relation to massed and distributed practice, was made by McDonald and Allen (1967). Four treatment groups were set up, and all students taught three 20-minute videotaped lessons, with the number of probing questions as the dependent variable. Practice and supervisory feedback were arranged as follows:

Group A - Feedback immediately following teaching, on three successive days

Group B - Feedback immediately following teaching, with a week's interval between each session

Group/

Group C - Lessons at weekly intervals, with feedback on each lesson delayed by a week (e.g. in week 2, students received feedback on their first lesson and immediately taught their second lesson)

Group D - Teaching and feedback in alternate weeks over a 6-week period.

No significant differences were found between groups, though a post-test after seven weeks indicated better retention of the skill by the delayed feedback and distributed practice groups. Turney (1973) has suggested that videotape replay has the effect of "reinstating" the student's performance for him, giving him the impression of immediacy even if feedback is delayed. As for the optimum interval between teach and reteach, the research evidence is inconclusive, though common sense would suggest that the interval should be long enough to allow the student reasonable time for replanning. The Minicourse booklets advocate a 24-hour period between teach and reteach. Levis (1973) has reported no significant difference in performance between students given a 20-minute break between teach and reteach and those given a week for revision and replanning.

Ward (1970), in his survey of American microteaching programmes, found that only a quarter used the complete teach-reteach cycle on all occasions and that a quarter omitted the reteach altogether. Brusling (1974) quotes a report by Skailand (1972) on the field testing of Minicourse 18 ("Teaching Reading as Decoding"). Skailand "found/

"found no effects of a number of variations in reteach treatment and concluded that reteaching may be valuable only when skills are the opposite of the skills usually in the teacher's repertoire". Orme (1966), in his work on modelling and feedback, also found some support for his hypothesis that initial gains would be greater than gains in later phases of training. These findings would appear, on the face of it, to clash with the conclusion of Allen, Fortune and Cooper (1968) that two complete cycles of skills training were needed in order to make microteaching effective.

(iv) Feedback : the supervisor's role

Introducing a concentrated period of microteaching into a busy college term creates many planning problems, not least the reallocation of tutors' time. A research study raises particular difficulties in this respect, since, in order to ensure the impartiality of assessments, no tutor involved with the experimental group can be allowed to evaluate the students' subsequent performance on teaching practice. It would therefore have been a convenience for the Jordanhill History department if research evidence had endorsed a decision to dispense with a supervisor altogether during the microteaching training sessions. Borg et al (1970) interpreted the weight of evidence in this way, citing studies by Acheson (1964), Orme (1966) and Claus (1968) to support their contention that supervisor comment at the feedback stage did not add significantly to the/

the performance gains derived from perceptual modelling and self-evaluation of videotaped lessons; and a study by Tuckman and Oliver (1968) which indicated that supervisor feedback alone led teachers to change their behaviour in a direction opposed to that recommended by the supervisors. The evidence for and against the value of supervision at the feedback stage has been carefully sifted by Griffiths (1972), in a review which draws attention to ambiguities which make it difficult to generalise from much of the research in this field. Borg, for example, considered that "the critical variables in Orme's (1966) study were perceptual modelling and videotape feedback; apparently, if these are present, supervisor feedback is unnecessary"; whereas McDonald and Allen (1967) interpreted their work on feedback variables, on their own and in relation to perceptual and symbolic modelling procedures, as a clear indication that "for producing some kinds of behaviour change, a modelling and feedback condition with an experimenter present during both phases is a powerful treatment". In fact, the writers went so far as to advocate a simple decision rule, that in mounting a microteaching programme one should "always include a feedback system in which the trainee views his own performance with supervision".

Perhaps the strongest argument in support of Borg's claims for self-evaluation as opposed to supervisory feedback is the impressive list of significant performance gains achieved during field tests of the Far West Laboratory's Minicourses (Borg et al, 1970). However, a study by

Morse/

Morse, Kysilka and Davis (1970) indicated that non-directive comment from a supervisor at the feedback stage was a significantly effective factor influencing subsequent use of "refocussing behaviours". The evidence of studies by Salomon and McDonald (1969) and by Tinsley and Davis (1969) is also relevant. Salomon and McDonald found that students viewing their own teaching performance without receiving any criticism did not achieve any "desirable attitudinal or behavioural changes"; while Tinsley and Davis found that student teachers who had been asked to plan questions on given material and assess the cognitive levels of their questioning produced self-ratings which showed very little significant correlation with the ratings of skilled observers. It would be reasonable to infer from this finding that, had these students been limited to self-evaluation of a skill such as higher order questioning, they might well have misinterpreted their own performance.

Commenting on the difficulty of drawing firm conclusions from research on the value of supervision in the micro-teaching process, Griffiths (1972) has suggested that a supervisor's effectiveness may depend on

- his role in relation to other components, such as modelling

- the student's level of competence in the skill to which the supervisor is drawing attention

- the expectancies which students have about the supervisor's/

supervisor's role

- the strategies of support, guidance and direction employed by the supervisor.

Evidence on the last of these factors is clearly of great concern to an institution embarking on microteaching for the first time. Acheson (1964), in a microteaching programme designed to reduce the percentage of teacher monologue and increase the frequency of pupil participation, investigated the effect of two types of supervisory treatment: direct supervision (suggesting changes for improvement in performance) and indirect supervision (eliciting suggested changes from the student). He also used a "no supervision" group as a control. No significant differences associated with supervisor-treatments were found for either of the behaviours. Johnson (1967) found some evidence that a supervisory approach which he described as inductive produced significant gains in the skill of giving directions; whereas a group receiving directive and methods-oriented supervision achieved more modest gains in the skill, and a group receiving directive and subject-oriented supervision achieved no gains at all. McDonald and Allen (1967), in their experiment to compare the effects of self-evaluation of performance with feedback from a supervisor, arranged that one group should view their lessons with a supervisor who simply reinforced (i.e. commented approvingly on) each use of the target skill, and that one group should view with a supervisor who gave discriminatory cues (i.e. pointed out instances when the use of the skill would be appropriate and suggested ways

in/

in which the skill might be used) as well as providing reinforcement. Two other groups viewed their lessons without a supervisor. It was found that the group receiving discrimination training and reinforcement from a supervisor did significantly better than the other three groups, which in fact regressed during the experiment.

Griffiths (1974) has reviewed a number of studies by Blumberg and various associates, related to his Categories for Analyzing Supervisor-Teacher Interaction, though not applied in a microteaching context. Evidence from these studies suggests that teachers react most favourably to indirect supervisor behaviour. These indications find support from the work of Ivey and Rollin (1974) in the field of microcounselling. A number of counselling skills were defined in behavioural terms and developed through an "each one - teach one" format, in which trainees first practised the skill and then taught it to other students. Evaluation centred on "a negotiation between the trainee and facilitator. The role of the supervisor was first to check that the trainee understood the nature of the skill and became reasonably proficient in it; and then to encourage the trainee to define his own level and style of performance" until he reached the point when he could "strike off in his own direction and operate independently from the trainer".

It might be argued that the expectancies of student teachers in relation to supervision would differ from those of the more mature personnel likely to be trained as counsellors. An investigation of student trainee role expectations of

the/

the microteaching supervisor was carried out by Johnson and Knaupp (1970). Analysis of a questionnaire given to 224 students indicated two dominant expectations: "first, that (the supervisor) be qualified to render technical assistance in planning for, conducting and evaluating instruction; second, that he give them an unhampered opportunity to find their own teaching style". The authors conclude that "certainly one might expect these students to resist, or even reject, a highly directive supervisory regime, at least until its merits were demonstrated".

There was also some indication from Johnson and Knaupp's study that students desired to share their microteaching experiences with their peers. The choice between group supervision and dyadic confrontation at the feedback stage was of concern to the Jordanhill project. Group supervision would clearly save time and concentrate the students' training experience. No firm evidence of the superiority of one or other treatment could be found in research reports, though a study by Young D A (1970) suggested that students working in teams with one of their number appointed as a supervisor performed significantly better in a number of skills than students receiving feedback singly from a regular supervisor. In the projects carried out at the University of Sydney by Turney (1970) and Owens and Hatton (1970), students asked for peer group assessment and subsequently expressed satisfaction with this approach. At the University of Stirling, McIntyre (1971) compared three treatments at the teaching and feedback stage of a microteaching programme offering training

in/

in three skills. One group worked individually with supervisors and practised all three skills. A second group worked in teams of three, each team with a supervisor, and each student practised one of the skills and observed, analysed and discussed the other two skills, practised by the fellow members of his team. A third group followed the same procedure as the second group, but worked without a supervisor. The first group showed significant superiority in the acquisition of the skills, but it was impossible to say, on the basis of this experiment, whether their superiority owed more to the individual guidance received from their supervisors or to their more extensive opportunities for practising their teaching in the microteaching context. There was no significant performance difference between the second and third groups, but most of the students in the experiment considered it advantageous to have a supervisor, and there was some evidence of low morale among the group that had worked without supervisory assistance. McIntyre's conclusion, which naturally carried weight with Jordanhill staff, was that "at least for these Scottish students without previous teaching experience, the provision of tutors appears necessary".

(v) Feedback : audio v. audio-visual recording

By comparison with video operation, audio tape recording at the feedback stage of the microteaching cycle can effect considerable savings of time; time for rigging equipment/

equipment, technicians' time, and time spent in locating extracts (since audio tape can be run backwards and forwards faster than video tape). Several microteaching reports, however, had expressed a belief in the superiority of video recording and playback. Unruh (1967) found that student teachers working form audio-visual recordings of lessons were able to make more accurate ratings of teacher effectiveness than students working from visual only, audio only or transcript records of the lessons. Gall (1971) noted that students preferred videotape to audiotape feedback, even though there was no overall significant difference in effectiveness between the two treatments. Wright (1973), in a study which employed both audio tape and video tape feedback, found a slight student preference for television recordings, while the tutors involved in the study rated audio feedback as less useful than either video feedback or feedback without the use of technical aids.

Common sense would support the superiority of video feedback for skills involving movement or gesture. The Jordanhill study, however, was concerned with verbal skills, and the evidence from previous studies which had compared the effectiveness of audio and video feedback in relation to such skills gave no consistent indication of superiority either way. Klingstedt (1970) compared verbal comment from peers and supervisor, unsupported by any recorded playback; comment supported by audio playback; and comment supported by video playback. No significant differences emerged from the three feedback treatments, even though the target skill was stimulus variation/

variation, involving both non-verbal and verbal behaviour. Acheson and Tucker (1971), in a training programme based on Minicourse 9 (Higher Cognitive Questioning), compared two instructional treatments:

- videotaped demonstration followed by videotaped practice

- videoscripted demonstration followed by audiotaped practice

Each group showed a gain of 62% in their use of higher order questions and the mean length of pupil responses in both groups increased by approximately equal amounts. Gall et al (1971) conducted a similar comparative study based on Minicourse 5 (Individualizing Instruction in Mathematics). Their overall conclusion was that audio and video feedback were equally effective and that the students' preference for video recording must be set off against the saving in cost and time when audio recording was used. In detail, videotape proved significantly more effective as an aid to training in demonstration techniques, but audiotape was superior as a feedback technique for training in methods of evaluating pupil progress. Bortz (1971) compared the usefulness of video and audio feedback in the context of training student teachers in higher order questioning and probing. No significant difference was found between treatments in respect of higher order questioning, but the audio group were slightly superior in their use of probing. Brusling (1974) tested the hypothesis of differential effects as between audio and video feedback in a programme to train

graduate/

graduate student teachers in the use of Flanders' Interaction Analysis Categories. Again, no significant differences were found between groups, though the audio group showed slight superiority.

Two studies produced rather more positive evidence in favour of audio feedback. Shiveley (1970) studied the effect of feedback modes on microteaching by comparing four groups of teachers

- videotape feedback viewed with supervisor

- audiotape feedback viewed with supervisor

- supervisor feedback based on his own observations of the lesson, without any recorded replay

- supervisor feedback based on Stanford Teacher Competence Appraisal Guide ratings, completed by pupils after the lesson.

Assessment of performance gains on the reteach lesson were based on pupils' S.T.C.A.G. ratings of the teachers. The audio group did best on all thirteen ratings of the S.T.C.A.G., and the teachers valued this mode of feedback highly. The video group showed little change, although video feedback was highly valued by the teachers concerned. Supervisor feedback based on pupils' ratings produced behavioural change, but the teachers did not value the approach highly. Supervisor feedback based on personal observation (in effect, the traditional form of teaching practice critique) was least effective and was not valued.

Shiveley/

Shiveley surmised that, in the context of verbal skill training (his report does not specify which skills were practised), video feedback may distract the student towards irrelevant aspects of his performance. Ward P M (1970) carried out a similar comparative study with four groups of elementary school teachers, using video feedback, audio feedback, a combination of video modelling and feedback, and reflective evaluation without replay of any kind. In relation to the skill of probing, he found that the audio group did best, and he suggested that audiotape recorders were "grossly under-rated" as a feedback device.

By contrast, Leonard et al (1971) found that video feedback proved more effective than audio feedback in training student teachers of English to move from direct to indirect styles of teaching. Both audio and video groups did better than a control group which received supervisor feedback without recorded replay, though the differences in magnitude of change among the three groups were not significant. The result is interesting, in that one would expect English specialists to be "verbalisers" rather than "visualisers", especially when they were concerned with the analysis and practice of predominantly verbal skills.

(f) The assessment of teaching performance

It will be recalled that one aim of the Jordanhill research study was to examine the extent of agreement between assessments of students made by lecturers using an appraisal instrument on school visits/

visits and assessments of the same students based on a detailed quantitative analysis of recorded classroom lessons. Guidance was therefore sought from previous research on the validity and reliability of traditional forms of teaching practice appraisal and on the comparative merits and limitations of high and low inference forms of assessment.

Wragg (1973), summarizing research studies which had attempted to establish criteria of effectiveness in relation to the assessment of student teachers, concluded that "there is comparatively little agreement amongst researchers about one single criterion of good teaching, and studies which have taken pupil gains on achievement tests, and pupil or observer ratings, have produced quite different groups of "good" teachers". Evidence to support Wragg's view is not difficult to find. Robertson (1957) drew up a list, obtained from teaching practice supervisors, of 50 "attributes of a successful teacher" and then asked 18 supervisors to rank the attitudes in order of importance. Correlations between rank orders ranged from +.73 to -.164, with an overall coefficient of concordance of +.377. Robertson concluded that the evidence "does not support the practice of awarding teaching marks or assessments without a valid scheme of reference". Shipman (1966) analysed 1,000 teaching practice assessments given in ten primary and ten secondary schools over a 5-year period, checking on the proportion of credit and distinction marks as between school and school and as between school and college. He found that the type of school to which a student was assigned had an important influence on the final teaching mark and concluded that, while grading of students enabled the best and the worst to be sorted out, "teaching practice marks do not seem to be a measure of true performance". In a study carried out at the University of Sheffield, Poppleton (1968) used

a/

a 27-item rating form to assess the ability of student teachers. School assessments based on this rating form achieved a +.60 correlation with those of university supervisors. A further analysis was made to determine, for each assessor, which of the 27 items correlated most highly with the overall teaching mark, on the assumption that the items with the highest correlation were those to which the assessor attached most importance in making his global rating. Among the school assessments, fifteen items achieved a positive correlation of .5 or over. Nine items achieved a .5 or over correlation among the university assessments, but only six of these items were subsumed in the fifteen school assessment items. The implication drawn by Poppleton from these results was that university supervisors attached importance to fewer aspects of teaching behaviour than did school assessors; and that there was an element of disagreement, as between university and schools, about which aspects of behaviour were most important.

One reason for the unreliability of supervisor assessments of teaching ability may be educed from the work of Cicirelli (1969). He made a content analysis of a large number of unstructured evaluations of student lessons, derived 23 categories of supervisor statements and coded a sample of the reports in terms of these categories. A calculation of the relative frequency with which each supervisor used each category enabled Cicirelli to produce a number of measures, showing the concentration of observers upon teacher-pupil relationships, the range of categories used by each observer in his assessments and the relative emphasis upon broad aspects or specific details of teaching behaviour. These measures were then correlated with the supervisors' scores on a test of creativity and the results indicated that the more creative supervisors took more notice of teacher-pupil relationships, used a wider range of categories/

categories and tended to emphasise broad, general aspects of behaviour.

A number of studies (for example, Wiseman and Start, 1965; Shipman, 1966) have indicated that supervisor assessments of students are unreliable in the sense of being poor predictors of future ratings of "success" in service. However, a study by Collins (1959) suggested that training assessments may be reasonably good predictors in making a broad distinction between "poor" and "good" teachers. 115 students were categorized, on the basis of their teaching marks, as either poor or good. Subsequent ratings made by the heads of the schools in which the students found jobs showed a correlation of +.57 with the supervisors' ratings and the difference between the school ratings of the two groups was highly significant. Collins estimated that 75% of the population achieved approximately what one would have expected from their teaching practice performance.

Since the Jordanhill History department, while insisting on the importance of high inference assessments of teaching ability, were prepared to adopt more detailed forms of appraisal than they had been using hitherto, it was useful to survey assessment procedures in other institutions. Stones and Morris (1972) analysed 122 responses to a questionnaire on approaches to teaching practice assessment, sent to colleges of education and university education departments. They found that the majority of institutions used a 5-point rating scale based on impressionistic methods of assessment, related to observation of a number of lessons. About half of the institutions completed a schedule of criteria of effectiveness, the three most commonly mentioned factors being teaching performance, planning and preparation of lessons and desirable traits in the student, in that order. No clear pattern emerged

from/

from the very wide variety of criteria in use and it proved difficult to group them under conceptual headings.

Rating forms as used by teaching practice supervisors were considered by Biddle (1964) to be "a plethora of hastily constructed and unreliable inventories". He maintained that "generally the results of research using rating forms have been poor and contradictory". Tittle and Hӓndle (1970) reported several studies (Anderson and Hunker, 1963; Lawler, 1964; Medley and Mitzel, 1959, 1963) which showed that supervisor ratings were not significantly correlated with pupil gains in the classes taught by the students concerned. Oppenheim (1970), in a general review of attitude measurement, pointed out that the correlation between raters is normally low and maintained that "the use of ratings invites the gravest dangers and possible errors, and in untutored hands the procedure is useless. Worse, it has a spurious air of accuracy which misleads the uninitiated into regarding the results as hard data".

Evidence of this kind must cast doubt upon the value of high inference assessments, which are necessarily based on some form of rating. Nevertheless, Rosenshine and Furst (1971), reviewing fifty studies which related observer teacher behaviours to pupil gains, found that the variables showing the most consistent correlation with pupil achievement (for example, clarity of presentation, variability of approach, enthusiasm and task orientation) were identified through high inference ratings of performance. The authors suggested that "the use of both high-inference and low-inference measures in future studies may be most advantageous. Rating scales may allow a student or an observer to process a large number of cues before he makes a decision on

a/

a teacher's "task-oriented" behaviour, whereas someone using an observational category system is unable to perform such processing because of the nature of the system. At the same time, the low-inference measures can provide specific details on "task-oriented" behaviour which might escape an investigator limited to high-inference measures".

Several writers have supported the complementary nature of high-inference and low-inference assessments of teaching ability. White (1972) stressed the importance of adapting teaching behaviour in a varied way to different goals, situations and responses, and argued that judgements on appropriate use of behaviours must be made separately from observations on the occurrence of the behaviours. Walker (1972) emphasized that context and content should be taken into account in making assessments of effectiveness. It was not sufficient merely to record behaviour. Parlett and Hamilton (1972) contended that assessment based exclusively on objective methods and quantitative measures was "artificial and restricted in scope". They considered that lessons, set in their social context, should be evaluated as a whole, through a process of observation, inquiry and explanation.

Finally, a recent study by Brusling (1974), aimed at training students to achieve a more indirect teaching style, provides a reminder of the importance of agreement on values and goals, if high and low inference assessments are to complement and not conflict with each other. Brusling found that, whereas his training programme implied a reduction in teacher talk, there was a positive correlation between amount of teacher talk and the grades on teaching competency awarded by classroom supervisors to students/

students taking part in the experiment. This finding is not, of course, generalizable to other institutions in different countries, but it is easy to envisage the confusion that might arise if, for instance, supervisors who set little store on the value of higher order questioning interpreted the "appropriate" use of such questioning as a very sparing employment of the skill. In order to avoid misunderstandings of this kind, the Jordanhill project spent a good deal of time during the first year on reaching agreement among tutors on the definition and desirability of the skills to be practised.

3. Method and procedure: the preparatory year

This process of reaching agreement characterised the research study throughout the first year, since the History department lecturers were required to play a dual role. In part, they were among the subjects of the study, but in part they helped to shape or at least approve its design. Between researcher and lecturers there was a continuing dialogue, to which the historians contributed their specialists' knowledge of the skills appropriate to history teaching at the secondary stage and advice on the art of the practicable, particularly as it related to the evaluation of student lessons. A good many of the weekly departmental staff meetings throughout the year were devoted to discussions of this kind, as well as to practical training in the use of appraisal instruments. Agreement at this level is achieved rather than born, and there was a number of second thoughts on content and design. In general, however, the work of the preparatory year followed a logical sequence, beginning with the identification and analysis of skills, leading to the preparation and testing of an Appraisal Guide, the selection of skills for micro-teaching practice, the recording of skills models and the formulation of a programme for the main study in the following session. At the same time, pilot microteaching practices were held, in the college and in neighbouring schools, to study logistical problems, assess the students' ability to interpret the analysis of selected skills and give tutors a chance to learn the techniques of supervision. Progress in each of the main stages of the work is described below.

(a) Identification and analysis of skills.

At the start of the project, the History department lecturers, new to the idea of microteaching and unused to thinking of teaching performance in terms of a comprehensive list of technical skills, were not prepared to make an immediate choice of skills for special study/

study and practice. Moreover, their first intention was to examine the effectiveness of microteaching as a remedial technique. ("Remedial" was used in a broad sense, to apply to training in any aspect of teaching in which a student was comparatively weak. It was not intended to restrict microteaching to students with serious difficulties.) It was therefore felt necessary to identify a comprehensive range of skills appropriate to all aspects of teaching behaviour. In addition, the lecturers wished, if possible, to relate performance on specific skills to teaching performance as a whole. For these reasons, it was decided to produce an appraisal instrument, broad enough to provide a reliable assessment of overall competence, but sufficiently detailed to give an indication of ability in each of the major skills.

The first task was to agree upon the concept of a "teaching skill". Morrison and McIntyre (1973) have defined it as "any pattern of teaching behaviour which generally tends to be effective in achieving a particular type of objective, and which a teacher uses sufficiently often for it to be a largely automatic response to the type of situation for which it is appropriate." This definition, had it been available at the start of the Jordanhill project, would have pin-pointed the main conceptual difficulties which arose during the first term of drafting and discussion. These related to

- the apparent overlap among skills, however defined and structured
- the breadth of the pattern of teaching behaviour which could be said to constitute a skill
- the question of whether one should limit the definition of "skill" to habitual, "automatic response" behaviour, or extend it to include more deliberate behaviour consciously applied in situations indicated by the teacher

For the History lecturers, these issues were not simply matters for theoretical/

theoretical debate. They had a direct bearing on training and assessment procedures. The time available for practical methods tuition in the twenty-eight weeks of the graduate course is necessarily brief; and in an attempt to lead students to a balanced competency, tutors tend to discuss the skills of teaching in broad terms, such as "effective questioning" and "clarity of exposition". Analysis of these broad patterns of behaviour revealed that each subsumed a number of sub-skills; but it was arguable that the detailed delineation of numerous sub-skills might confuse the students and leave the tutors, in their assessment of student lessons, groping among the trees without any clear view of the wood. Again, the problem has been well stated by Morrison and McIntyre (1973):

> "One question is how narrowly to define each skill. It seems probable that skills defined in terms of only one or two behavioural criteria can be more effectively mastered in microteaching; but they may not be assimilated any more easily into a student's teaching repertoire than more broadly defined skills, on the practice of which it is possible to spend more time."

In an attempt to meet training needs (though perhaps at the expense of consistency) the Jordanhill study distinguished three categories of teaching skills:

(a) basic: related to habitual patterns of teaching behaviour, such as exposition, questioning and reinforcement

(b) composite: skills such as varying the pace and rhythm of a lesson, which in fact result from blending a number of basic skills - questioning, varying the stimulus, reinforcement, etc. - in proportions which give coherence and variety to the lesson as a whole

(c) specialist: related to the teaching of a particular subject; for example, the demonstration of scientific experiments, the/

> the presentation of historical documents. Such skills often entail conscious planning and applications by the teacher and they may involve the use of basic sub-skills.

Within these categories, eighteen broad skills (or, to use the Stanford phrase, "major teacher competences") were identified. They were intended to cover the whole range of teaching behaviour, at the planning stage (five skills) and the presentation stage (thirteen skills). Each broad skill was analysed into a number of sub-skills, described in behavioural terms, to facilitate application by the student and assessment by the tutor. This analysis resulted from a series of drafts, exchanged between the researcher and the principal and senior lecturers of the history department and subsequently discussed by the department as a whole. Care was taken to word the descriptions in a way that would be appropriate to the disciplines of history teaching. For example, an early gloss, submitted by the researcher, on Coherence of Presentation stated that "the teacher presented the subject matter in a logical sequence." The senior lecturer objected that historical events could not be said to occur in a logical sequence, and he therefore substituted "orderly and intelligible sequence". The importance of agreeing upon the analysis of each broad skill, as a prerequisite of reliability in assessing teaching competence, was acknowledged by all the lecturers. Although the initial list was established after some two months' work, it was revised once more as the year progressed, in the light of reflection and experience, and re-ordered under the following headings:

Preparation/

Preparation
- Objectives
- Choice of Content
- Planning of Method

Presentation
- Clarity and Coherence
- Content
- Stimulus
- Use of Learning Aids
- Varying the Pace and Rhythm
- Evaluation
- Books and Documents

Pupil Involvement
- Question Technique
- Pupil Participation
- Organization

Relationships
- Relationships (learning atmosphere, efficient control, friendly rapport)

The analysis of each of these fourteen broad competences (set out in Appendix A) was used as the basis of the main experiment in the second year.

(b) Preparation and testing of Appraisal Guide

The Appraisal Guide in its first form (see Appendix B) consisted of a list of the eighteen skills originally identified. Each skill was glossed by a descriptive sentence summarizing the main points of the analysis. Assessments were made on a 7-point rating scale. A rating procedure was adopted, in spite of the fact that some researchers (e.g. Biddle, 1964; Oppenheim, 1970) had questioned the reliability of rating scales and instruments, since the skills analysis incorporated the concepts of "appropriate use" and "taking advantage of opportunities", thus involving the assessor in high-inference judgements which implied rating rather than quantitative measurement.

Early experience of using the Appraisal Guide produced a crop of problems, both conceptual and practical, which were examined in a series of/

of staff meetings. Achieving inter-tutor reliability involved agreeing on a concept of "average performance", but discrepancies in the assessment of videotaped student lessons revealed a number of different approaches to performance norms. Some tutors based their assessment on "what could reasonably be expected from students at this stage of their training" and thus applied increasingly critical standards as the year progressed. Some tutors took "average" to mean "average performance based on my total experience of student lessons". Some related their assessment to a generalised concept of "an averagely effective lesson taught on this particular subject to pupils of this age and stage". Some tutors, more than others, took account of pupil reaction as well as teacher performance. Some, but not all, made allowances for difficult classes and unfavourable school environments.

There were also varying approaches to the range of assessments given. Some tutors attempted to relate the points in the rating scale to absolute standards of performance. Thus, an "A" was associated in their minds with a "near perfect" lesson and was rarely if ever awarded. Other tutors thought of the rating scale in percentile terms, although no general decision had been taken about the proportion of the population covered by each point on the scale, and accordingly tended to award a broader range of assessments.

Problems of weighting affected the reliability of overall assessments. Tutors found it difficult to agree about which skills were "most important". For example, group assessments (and subsequent discussion) of videotaped lessons indicated that some tutors attached particular importance to ensuring pupil participation, while others put more emphasis on presentation skills. Thus, two tutors might agree in their ratings of these specific skills, but still differ widely in their overall/

overall assessment.

Morrison and McIntyre (1973) note that "while one may be confident that a specified skill is of value in teaching, the decision as to whether or not it is appropriate to use that skill in the context of any particular lesson must generally be highly subjective." It will be remembered that the Jordanhill tutors attached great importance to appropriateness and making the most of opportunities in the use of a skill. It would be expected that the element of high inference thereby introduced into judgements of performance would tend towards unreliability, even when tutors were in general agreement about the kind of teaching behaviour which was desirable in a particular context. Analysis of student lesson ratings revealed, however, that such agreement did not always exist. For example, tutors showed marked disagreement (ranging from 2 to 7 on the rating scale) in assessing the skills relating to Pupil Involvement in a lesson taught by a mature student to a first year class of low ability. The subject-matter of the lesson had been clearly presented, but questions had been restricted to a cognitive level demanding only brief, factual respnses. In the course of discussion, some tutors asserted that pupils of this type were not capable of reasoned responses and that attempts to involve them at a higher cognitive level would therefore have been "inappropriate". Other tutors were equally convinced that such pupils could not cope with long periods of passive listening to exposition (however well presented) and that the student had missed many opportunities to involve the pupils in active response. Disagreements of this kind, reflecting differences of attitude, personality and experience, remained unresolved throughout the session.

In observing student lessons, tutors saw themselves as filling a number of roles: diagnostician, counsellor, critic and assessor. Some tutors associated/

associated the use of the Appraisal Guide with assessment rather than diagnosis, and were consequently concerned about confidentiality. Had the student the right to see this detailed report? What would happen if he discovered that he had been given low ratings? Some effort was required to dispel this kind of anxiety, which may well have led some tutors to skew their ratings towards the upper end of the scale.

Practical problems in using the Appraisal Guide were eased in the latter part of the year, when the Guide was reconstructed in a simpler form (see Appendix C), requiring tutors to make only four assessments of aspects of the lesson as a whole, but calling for more detailed assessment of performance in those skills which were to be practised in microteaching sessions. This revision, however, failed to solve all the tutors' problems. They found difficulty in deciding upon an appropriate rating in those instances where a student made little use of a group of skills (Pupil Involvement, for example), not because he had failed to grasp his opportunities but because the frequent use of the skill was not appropriate to that particular lesson. There were problems, too, in rating "avoidance behaviours" (for example, "avoiding over-use of one-word answer type questions"), in lessons where a student avoided such over-use simply by asking very few questions of any kind. A different type of problem arose for those tutors who were accustomed, in making their assessments, to attach great importance to "teaching personality". The Analysis of Skills had deliberately made no direct reference to personality traits (confidence, impatience, sense of humour, etc.). Its definitions were expressed in behavioural terms, since it was felt that behaviour could be more objectively observed (and therefore more reliably assessed) than personality as such. This behavioural approach proved readily acceptable to some tutors, but others felt that the Appraisal Guide omitted important factors which they wished to take into account in making their/

their assessments.

The most intractable problem for the majority of tutors was that of combining assessment of specific skills with judgment of the general quality of a student's teaching. To appreciate this quality, they felt that they needed to sit back from the lesson, to be aware of its overall rhythm and balance, to sense its cummulative impact upon the pupils. They accepted that the ability to combine a wide and narrow focus of observation would probably grow with training; but the time which they could spare for such training was limited, and meantime they were concerned lest, in concentrating upon the pulse tick of analysis, they should miss the heart beat of the lesson as a whole.

At this early stage of the research study, the plan was to rely solely on tutors' assessments in analysing student progress during the main experiment in 1972-73. It was therefore essential to demonstrate that the Appraisal Guide could be reliably used by the history lecturers. It would have been helpful to make direct measurements of intra-tutor reliability, since students on teaching practice were visited two or three times by the same tutor in the course of a half term out in a school; but this would have involved repeated assessments of the same lessons, with long intervals between trials, to avoid the danger of achieving a false consistency due to recall of previous ratings. Within the limits of the preparatory year, such spacing was impossible. However, it was felt that, if inter-tutor reliability could be demonstrated over a range of lessons, intra-tutor reliability could reasonably be inferred, since repeated agreement among tutors would indicate that they were applying consistent criteria in making their assessments. It was necessary, in any case, to demonstrate inter-tutor reliability, as a basis for comparing the performance of control and experimental groups and for measuring Term 1 to Term 3 performance changes./

changes. (As a matter of policy, the History department allocated fresh tutors to students at each stage of their teaching practice.) Such reliability should cover both overall lesson assessments and ratings for specific skills and sub-skills.

It was accepted that, since the Appraisal Guide involved an element of subjective judgment, complete reliability would rarely if ever be achieved. In the long term, if the department were to continue to employ microteaching as a remedial technique, the most important consideration was reliability of decision-making; for example, agreement on the weakest elements among a student's teaching skills. In the short term, for research purposes, it was important, if the Appraisal Guides were to be used to measure and compare progress, to achieve a high degree of reliability in terms of scores on the rating scales. Accordingly, in analysing the evidence of the reliability trials, it was decided to examine both variance and correlation.

Four reliability trials were mounted in the course of the preparatory year. The first two, held in November 1971 and February 1972, employed the initial form of the Appraisal Guide. In the February trial, tutors attempted to assess the skill of "clarity and coherence" in some detail, in addition to completing the Appraisal Guide as a whole. The trials were based on videotaped student lessons, scored by all tutors in the department. Examination of the results revealed such wide variance among the tutors that there was clearly no point in making a full analysis until further training had been carried out. The two final trials, held in the following June and September, employed the revised form of the Appraisal Guide. The June trial was based on eight specially recorded student lessons. Since the appraisal was related to videotape playbacks, it was not possible to assess Preparation skills, but each lesson was scored for Presentation, Pupil Involvement and Relationships, and, in addition, detailed assessments were/

were made of the following skills:

Clarity and Coherence	-	2 lessons
Question Techniques	-	2 lessons
Using Documents	-	2 lessons
Varying the Pace and Rhythm	-	2 lessons

These were the skills on which it was hoped to concentrate throughout the main experiment. The students concerned had been asked to study the relevant skills analyses before teaching their lessons.

The September trial was also based on eight recorded student lessons, but assessments of the first two lessons were anslysed and discussed on a group basis by the tutors, in a final attempt to resolve disagreement, and statistical procedures were confined to the ratings of the remaining six lessons. By this time, the design of the main experiment had been changed (see Chapter 4) and accordingly all lessons were assessed on details of Question Technique, as well as being rated under the headings of Presentation, Pupil Involvement and Relationships.

The main statistic derived from the June and September assessments was a Coefficient of Reliability (r), using the formula

$$r = 1 - \frac{\text{error variance}}{\text{observed variance}}$$

where error variance = the variance among ratings made by tutors on each aspect of each skill for each student

and observed variance = the variance among all ratings made by tutors for all students on all aspects of all skills being analysed

Error variance was calculated from the formula

Ev/

$$EV = \frac{\Sigma x^2}{n} - \left(\frac{\Sigma x}{n}\right)^2$$

and observed variance from the formula

$$OV = \frac{\Sigma fx^2}{n} - \left(\frac{\Sigma fx}{n}\right)^2$$

where x = the rating awarded (from 1 to 7) by each assessor

n = the number of assessments

Additionally, in the June trial, coefficients of reliability were determined in respect of the varying extent and direction of differences in rating the selected skills on which student pairs were concentrating. Two sets of calculations were made: one concerned with variance of total differences for each pair of students and the other with the variance of differences for each aspect of each skill. The dame formula was used for calculating r, but, for variance of total differences

error variance - variance among total differences for each pair of students (9 differences)

observed variance = the total of variance over 36 total differences (i.e. 4 student pairs x 9 tutor assessments)

and, for variance of differences for each skills aspect,

error variance = variance among differences for each aspect of each skill (9 differences in each case)

observed variance = variance among all differences for all aspects of all skills (30 aspects x 9 tutor assessments)

The coefficients of reliability obtained from the June and September trials are set out in Tables I - IV.

These measurements of reliability were applied to the detailed assessment of the skills selected for microteaching practice. Reliability in appraising lesson performance as a whole was examined by a series of correlational tests, as follows:

(i) for each of the general teaching competences (Presentation, Pupil

Involvement/

Involvement and Relationships), the ranking of students by each tutor was correlated with the overall rank order (based on the sum of marks awarded to each student by all tutors). In view of the large number of tied rankings, the formula used was

$$\rho = \frac{\sum x^2 + \sum y^2 - \sum d^2}{2\sqrt{\sum x^2 \times \sum y^2}}$$

(see SIEGEL S, Nonparametric Statistics for the Behavioural Sciences, p. 207).

with an adjustment, $\sum T$, for tied rankings, based on the formula

$$T = \frac{t^3 - t}{12}$$

where t = the number of observations tied at a given rank.

(ii) a coefficient of concordance was determined for the tutors' ranking of students in terms of their total lesson scores. The formula used was

$$W = \frac{S}{\frac{1}{12} k^2 (N^s - N) - k\sum T}$$

where S = sum of squares of the observed deviation from the mean of Rj (Rj being the sum of ranks in each column of a K x N table)

k = number of sets of rankings

N = number of individuals ranked

$\sum T$ = corection for tied rankings

(see SIEGEL, p. 234)

TABLE I

Coefficients of reliability for assessment of specific skills

June, 1972

Clarity and Coherence	Student 1	Student 2
Clear voice	.61	.39
Suitable vocabulary	.57	.93
Examples to clarify points	.76	.77
Words explained	.21	.85
Historical concepts explained	.61	.79
Orderly sequence	.81	.81
Intelligible pattern	.52	.69
Key points stressed	.45	.76
Responsive to pupil difficulties	.56	.93
Question Technique	**Student 3**	**Student 4**
Frequent questions	.94	.52
Clear phrasing	.92	.66
Avoiding one-word answers	.70	.77
Avoiding repetition of answers	.67	.51
Varied levels of difficulty	.78	.59
Adequate time for response	.92	.64
Approval and tolerance	.77	.84
Constructive use of answers	.89	.70
Varied categories of questions	.78	.58
Using Documents	**Student 5**	**Student 6**
Documents put in context	.79	.61
Words, etc. explained	.70	.81
Interesting presentation	.92	.70
Drawing conclusions	.74	.54
Varying the Pace and Rhythm	**Student 7**	**Student 8**
Keeping balance	.64	.54
Varied pace and style	.18	.56
Varied rate of facts	.56	.41
Range of resources	.82	.04
Questions and discussion	-.09	.22
Pupil activity	.23	.12
Sections balanced	.54	.76
Variation overall	.54	.51

Average coefficient of reliability (all students, all skills) = .63

TABLE II

Coefficients of reliability for variance in total differences of assessment for student pairs:

June 1972

Student pairs	Coefficient of reliability
Student 1 - Student 2	-.06
Student 3 - Student 4	.32
Student 5 - Student 6	.85
Student 7 - Student 8	.52

TABLE III

Coefficients of reliability for variance in differences of assessment of each aspect of specific skills practised by student pairs: June 1972

Clarity and Coherence	Student 1 - Student 2
Clear voice	.56
Suitable vocabulary	.35
Examples to clarify points	.56
Words explained	.07
Historical concepts explained	.54
Orderly sequence	.23
Intelligible pattern	-.19
Key points stressed	-.20
Responsive to pupil difficulties	.33
Question Technique	**Student 3 - Student 4**
Frequent questions	.06
Clear phrasing	.62
Avoiding one-word asnwers	.46
Avoiding repetition of answers	.46
Varied levels of difficulty	.40
Adequate time for response	.59
Approval and tolerance	.53
Constructive use of answers	.52
Varied categories of questions	.11
Using Documents	**Student 5 - Student 6**
Documents put in context	.04
Words, etc. explained	.20
Interesting presentation	.67
Drawing conclusions	.76
Varying the Pace and Rhythm	**Student 7 - Student 8**
Keeping balance	.49
Varied pace and style	-.67
Varied rate of facts	.41
Range of resources	.09
Questions and discussion	-.67
Pupil activity	.40
Sections balanced	.59
Variation overall	.43

TABLE IV

Coefficients of reliability for assessment of specific aspects of questioning skills: September 1972

(n.b. Lessons by students 1 and 2 were used for discussion)

Questioning skills	Students						Skills averages
	3	4	5	6	7	8	
Avoiding one-word answer questions	.82	.84	.28	.75	.78	.25	.62
Adequate time given for response	.93	.78	.58	.69	.87	.57	.74
Making constructive use of answers	.91	.86	.35	.53	.82	.70	.70
Varying level and nature of questions	.86	.87	.53	.61	.84	.71	.74
Student averages	.88	.84	.44	.65	.83	.56	

Average coefficient of reliability (all students, all skills) = .70

(iii) the consistency of relationship among tutors in terms of the leniency/severity of their assessments was expressed by a coefficient of concordance based on tutor rank orders as shown by the overall scores given to each student. (e.g. Student 1 might receive his highest score from tutor B, his next highest from tutor F, etc; while Student 2 might receive his highest score from tutor C, his next highest from tutor D, etc. High correlation among these rankings would indicate a consistent relationship among tutors in terms of the leniency/severity of their assessments).

(iv) in the June trial, a correlational test was applied to tutor rank orders as shown by their assessments of the specific skills practised by student pairs.

The results of these correlational tests are set out in Tables V-VIII.

TABLE Va

Correlations between overall ranking of students on general teaching skills and ranking by individual tutors: June 1972

Tutor	Presentation	Pupil Involvement	Relationships
A	.94	.96	.81
B	.55	.82	.92
C	.98	.96	.87
D	.95	.94	.95
E	.95	.93	.74
F	.88	.87	.94
G	.94	.69	.81
H	.93	.81	.88
I	.70	.72	.82
Averages	.87	.86	.86

TABLE Vb

Correlations between overall ranking of students on general teaching skills and ranking by individual tutors: September 1972

Tutor	Presentation	Pupil Involvement	Relationships
A	.82	.74	.91
B	.68	.85	.76
C	.85	.91	1.0
D	.88	.94	.58
F	.94	.94	.94
G	.88	.76	.76
I	.94	.74	.85
J	.88	.97	.85
K	.76	.74	.88
Averages	.85	.84	.84

(n.b. Two tutors, J and K, joined the research study between June and September. Tutors E and H did not score all tapes in September trial)

TABLE VIa

Coefficient of concordance for tutors' ranking of students on total lesson scores: June 1972

	Ranking by Tutors								
Students	A	B	C	D	E	F	G	H	I
1	6	3.5	5.5	5.5	7	4.5	7.5	7	8
2	7	8	7	7.5	6	8	5	5.5	7
3	1	1	2	1	2.5	1	1	1	1
4	4.5	3.5	3	3.5	4	3	3	3	3
5	4.5	6.5	5.5	5.5	5	6.5	6	5.5	6
6	8	6.5	8	7.5	8	6.5	7.5	8	4
7	2	5	1	2	1	2	3	2	2
8	3	2	4	3.5	2.5	4.5	3	4	5

Coefficient of Concordance = .82

TABLE VIb

Coefficient of concordance for tutors' ranking of students on total lesson scores: September 1972

	Ranking by Tutors								
Students	A	B	C	D	F	G	I	J	K
3	5	5	6	5.5	5.5	6	6	5	6
4	1.5	1	1	1	1	1	1.5	1.5	1
5	1.5	2	2	3.5	3	3	3	1.5	2
6	3	4	4	3.5	4	5	4	4	3.5
7	6	6	5	5.5	5.5	4	5	6	5
8	4	3	3	2	2	2	1.5	3	3.5

Coefficient of Concordance = .87

TABLE VIIa

Coefficient of concordance for tutor rank orders in respect of overall scores awarded to each student: June 1972

	Students							
Tutors	1	2	3	4	5	6	7	8
A	2.5	3.5	4.5	8	2.5	6	5.5	3
B	2.5	9	8.5	9	7	3	9	5.5
C	6	8	8.5	3	7	9	1.5	5.5
D	6	6.5	4.5	7	7	6	7	7.5
E	6	1.5	6.5	5	1	6	1.5	1
F	1	6.5	2	5	5	2	5.5	9
G	8	5	2	5	9	6	8	3
H	4	1.5	2	1.5	4	6	3.5	7.5
I	9	3.5	6.5	1.5	2.5	1	3.5	3

Coefficient of Concordance = .22

TABLE VIIb

Coefficient of concordance for tutor rank orders in respect of overall scores awarded to each student: Spetember 1972

	Students					
Tutors	3	4	5	6	7	8
A	5	3	1.5	1	8	6.5
B	2	8.5	6	6.5	6	8
C	8	3	3	4	6	5
D	5	3	9	9	6	4
F	9	3	4.5	2	9	1.5
G	5	6	8	8	1	6.5
I	5	8.5	7	4	3.5	3
J	1	3	1.5	4	3.5	1.5
K	5	7	4.5	6.5	2	9

Coefficient of Concordance = .20

TABLE VIII

Correlations between tutor rank orders in respect of scores awarded to student pairs for specific skills: June 1972

	Student pairs							
Tutor	1	2	3	4	5	6	7	8
A	3	6	6	9	5	5.5	4.5	2
B	5	7	8	7	5	2	9	5
C	1	8	7	3	5	8	1	3
D	4	9	2	3	4	3	7	9
E	6	1.5	4	3	1	5.5	2	1
F	2	5	4	6	8	4	4.5	6
G	9	1.5	4	8	3	8	8	8
H	7	3	1	5	9	8	6	7
I	8	4	9	1	2	1	3	4
Correlations	-.72		-.14		.24		.70	

The failure of the history lecturers to achieve a consistently high degree of reliability, after a year's experience of the Appraisal Guide and the Skills Analysis on which it was based, supports the findings of researchers such as Robertson (1957), Shipman (1966) and Cicirelli (1969), summarized in Chapter 2, Section f. Nevertheless, it must be emphasized that considerable progress was made between the trials of November/February and those of June/September, and that the training in the intervening period was by no means intensive, not through any lack of good will on the part of the tutors but because of existing heavy commitments.

Considered in more detail, the results in Table I show wide fluctuations of reliability, in assessments both of students and of skills aspects. In many instances a high reliability factor for performance in a particular skills aspect of one student is offset by a low factor in respect of the other member of the pair. It is noticeable that while the average reliability coefficients for the first three skills are reasonably consistent ($r = .67, .74, .73$), there is a marked drop ($r = .41$) for the composite skill of Varying Pace and Rhythm. It was to be expected that this type of extended skill, which could be evaluated only over the lesson as a whole, would be as tricky to assess as it was difficult to define.

Tables II and III indicate wide inter-tutor variations in the extent and direction of differences in the rating of student performance on common skills, both overall and in detail. The average reliability coefficients for judgment of differences compare badly with the corresponding figures for the assessment of specific skills:

Skills/

	Skills assessment	Difference judgment
Clarity and Coherence	.67	.25
Question Technique	.73	.42
Using Documents	.74	.42
Varying the Pace and Rhythm	.41	.13

High levels of reliability for variance in differences of student pair assessments would have indicated that, while may have differed among themselves in their assessments of any one student, these differences remained relatively constant from one student to the next. This, in turn, would have implied common criteria for judgment (though applied with varying degrees of leniency/severity) and, for each tutor, consistency of standards as between assessments. The low coefficients obtained for most of the skills categories mean that little reliance can be placed either on agreements or on disagreements among tutors. For example, the detailed difference scores contained a number of instances where tutors x and y had given student m the same rating for a specific skill; but the ratings given to student n for practising the same skill were, in the case of tutor x, higher than for student m and, in the case of tutor y, lower than for student m. The implications must be that the tutors were employing different criteria (which happened to produce similar ratings for one student, but which led to varying ratings for his partner); or that tutors were unstable in their application of agreed criteria from one assessment to the next; or that both these factors were contributing to unreliability.

The figures in Table IV show no real advance over those in Table I; and indeed the lack of improvement is disappointing, considering that only four detailed skills aspects were assessed in September, as against thirty in June. The results suggest that tutors found it much/

much easier to agree about some students than about others. This supports the evidence of the discussion sessions, that the juxtaposition of certain teaching styles and pupil groups (e.g. an authoritarian teacher with a class of low ability) provokes widely differing reactions among training staff.

Tables Va, b and VIa, b show that the tutors achieved consistently high reliability in ranking students in terms of broad teaching competences and overall performance. The figures support the department's contention that there was usually "a fair measure of agreement" when a student was rated in general terms by more than one tutor. Reliability over rank order would be helpful in identifying students who should be offered microteaching as a remedial technique, but it is not in itself a sufficient basis for assessments in a study concerned with measuring and comparing progress.

The low correlation coefficients shown in Tables VIIa, b and VIII demonstrate a lack of any consistent tendency for some tutors to be more lenient, or more severe, than others in their assessments. Variance among tutors in rating students on skills performance cannot, therefore, be attributed in any significant degree to a consistent leniency/severity pattern.

In general, the results of these two trials indicated that the problems revealed in practice sessions had not been fully resolved. Agreement on ranking students on general teaching performance was consistently high; but, while the overall level of reliability for assessment of specific skills appeared reasonably satisfactory, the wide variations in extent and direction of differences in judging student pairs indicated that it would be unwise to place much reliance on tutor ratings as the sole or even the major basis for assessment

in/

in the course of the main experiment.

Apart from the negative evidence relating to leniency/severity patterns, the results provided no certain pointers to the reasons for tutors' difficulties in achieving reliability; though there were indications that problems of agreement could relate to particular students (e.g. student 4 in Table I and student 5 in Table IV), and to the kind of extended skill categories characteristic of Varying the Pace and Rhythm. In planning future measures to raise the level of reliability of tutor ratings, the researcher was therefore forced back on notes of departmental discussions, comments included on completed Appraisal Guide forms and his own observation of the lessons which had been used for rating purposes. This evidence suggested that tutors were still finding difficulty in coping with the detail of the Appraisal Guide; that there were varying concepts of what constituted "average performance" and "appropriate use of skills"; that, in spite of apparent acceptance of the comparatively objective approach to observation implicit in the Analysis of Skills, a tutor might retain a private interpretation of the nature of a skill, with the result that he might not always assess the behaviour which the Appraisal Guide asked him to assess; and that, even within the relatively precise definition of each skill category, some tutors were placing special weight on particular details. (For example, there was a range of reactions to the occurrence of the glottal stop, which may have affected judgments on clarity of speech; and some tutors took particular exception to certain forms of words employed in phrasing questions.)

(e) Selection of skills for microteaching practice and planning of main experiment

The principal concern of the History department was that the skills selected/

selected for microteaching practice should be

- basic to the needs of beginning teachers
- sufficiently straightforward to allow students to acquire them readily at the pre-service stage of training
- skills which were known to present difficulties for at least a proportion of the students.

Conscious of the amount of ground which it was felt necessary to cover in the brief graduate training period, the department was alert to the danger of spending an unjustifiable amount of time practising small details of behaviour (for example, aspects of non-verbal reinforcement) at the expense of grasping broader skill patterns. Just as a driving instructor has to teach a pupil road sense as well as mechanical techniques such as gear-changing, so the lecturers saw their job as extending beyond basic skills training to providing insights into the appropriate selection of skills in different teaching contexts. They wished to discover if microteaching could help to develop such insights through a selection of "broad-band" skills. They also wished to experiment with training in different types of skills, basic, specialist and composite. Finally, they were anxious not to place undue emphasis on a "traditional skills" approach to training. During the 1971-72 session, the department was hoping to persuade schools to co-operate in allowing students to undertake group project work during their teaching practice in the summer term 1973. Accordingly, it was suggested that two "management of learning" skills should be included in that term's microteaching programme.

The History department was accustomed to training approximately sixty graduate students a year, equally devided between honours and ordinary graduates. The majority of these students spent terms 1 and 3 under the guidance of the department and studied a second subject in term 2. A minority of the students, who were offering History as their sole

teaching/

teaching subject, remained with the department throughout term 2 and they were joined by other graduates who were offering History as a second subject. Thus, in operating the half-term in college/half-term in schools system, the department had about thirty students on teaching practice at any time during terms 1 and 3, and about twenty students in term 2. In planning the design of the main experiment, it was decided at first

- to provide microteaching for ten students in each half of term 1; for five students in each half of term 2; and for eight students in each half of term 3: a total of 46 students

- to restrict the microteaching practice to two days for each student in the mid-week of teaching practice, thus allowing two complete training cycles per student

- to compare pre- and post-microteaching performance in the classroom, on the basis of tutors' assessments during the early and later weeks of each teaching practice period, and to set the results against the performance of control groups.

The skills originally selected for microteaching were

Term 1 - 1st and 2nd half - Clarity and coherence
Term 1 - 1st and 2nd half - Question technique

Term 2 - 1st and 2nd half - Using historical documents

Term 3 - 1st and 2nd half - Organizing group and individual work
Term 3 - 1st and 2nd half - Quiding pupil activity in group work

The two group work skills, adapted from the broad skills of Organization and Pupil Participation (see Appendix A), were clearly not suited to practice under normal microteaching conditions. By their nature, they demanded a complete class and an extended teaching/learning period. They also involved the planning and development of a project on which group and individual work could be based and this could not be done at a moment's notice with a class brought into the College from a nearby school, nor would students have time or facilities to develop projects specifically for microteaching purposes. The best solution seemed to be

to/

to record the students out in schools, at work on the projects which (it was hoped) would form a normal part of their term 3 teaching practice; to play back and discuss the recordings; and to assess the students' performance on the skills at a later stage of their projects. The practical problems of recording in a large number of schools over a short period would be considerable and it was for this reason that the experimental group was limited to eight students in each half of term 3.

There would have been great interest in assessing the effectiveness of this combination of skills analysis, modelling and feedback, without the elements of control, sealing down and repeated practice normally associated with microteaching. Unfortunately, towards the end of the session, the department had to abandon its plan for including project work as a regular part of teaching practice, since too few schools were prepared to co-operate. The two group work skills were therefore replaced by a new skill (developed from the skill of Pacing, included in the original draft of the Analysis) entitled Varying the Pace and Rhythm. Tutors were agreed that this composite skill was one which graduate students, after several years of learning from lectures, found particularly difficult to master. There was agreement, too, that this skill was among those which lay at the heart of effective teaching, but the problem of defining it in behavioural terms was acknowledged. It was felt to be one of those blends of behaviour, immediately recognizable by its absence, yet difficult to assess objectively and therefore not an ideal choice for a research study. The prior consideration, however, remained the needs of the students and, on these grounds, Varying the Pace and Rhythm seemed an obvious choice.

The final selection of skills for microteaching was therefore

Term/

Term 1 - Clarity and coherence
- Question technique

Term 2 - Varying the pace and rhythm

Term 3 - Using historical documents

It was felt that this selection would enable the department to test the effectiveness of microteaching across a broad sample of basic, specialist and composite skills. Interpretative notes (see Appendix D) on all the skills, to support the Analysis, were prepared for both tutors and students, and arrangements were made with local schools to supply pupils for microteaching sessions at the College throughout 1972-73.

(d) Pilot microteaching sessions and the training of supervisors

Microteaching sessions based on the selected skills were held each term throughout the preparatory year, with the three aims of

- studying the practical problems involved
- discovering the reactions of students and assessing their difficulties in tackling this new training technique
- giving supervisors experience in evaluating students' performance on specific skills, using a Teaching Skills Analysis sheet which was developed from the Appraisal Guide (see Appendix E)

These sessions were mounted both in College and in neighbouring schools. Neither arrangement presented any great practical problems, though on balance it appeared preferable to organize extended sessions (i.e. those lasting several days) in the college. This involved the expense of transporting pupils from their schools, but it allowed more time for setting up equipment beforehand, thus enabling more care to be taken in the placing of cameras and microphones. Tutorial staff naturally preferred to have the sessions in college, but so did the students, who were brought in from teaching practice. They knew their way to the college/

college, but they sometimes found difficulty in locating the schools where microteaching was being held. The indications from these pilot sessions were that, in future, two different systems should be devised for microteaching: a college system, based on trolley-mounted equipment, and a school system, based either on simple audio equipment or on a hand-held camera and tape recorder, which lecturers could transport, assemble and use without assistance, on their visits to observe students teaching in schools.

All students involved in these pilot sessions were given a brief explanation of the nature and purpose of microteaching and encouraged to study the analysis of the skill which they were being asked to practise. In spite of these preparations, a number of the students found difficulty in concentrating their ten minutes of teaching upon a specific skill and in covering the various aspects of broad-ranging skills such as Clarity and Coherence or Question Technique. It seemed clear that most students needed a more extended introduction to the concepts involved, supported by some kind of model of the skills.

In terms 1 and 2, practice in supervising these microteaching sessions was given to as many tutors as possible. By terms 3, it was decided that the design of the main experiment demanded a small "microteaching team". For one thing, it was highly desirable, in order to avoid bias, that tutors assessing classroom performance should not know which students had received microteaching. For another, it was felt that concentrated training for a small team of interested tutors would produce a better quality of supervision. Three tutors (supported by the senior lecturer acting as standby) were therefore asked to meet for weekly training sessions throughout the summer term. The plan was that, during the following session, these tutors should concentrate on micro-teaching, leaving school visiting to be handled by the other seven tutors/

tutors. Eight sessions were arranged, covering the four selected skills. For each skill, tutors viewed, assessed and discussed two tapes in week A, and viewed and assessed a further four tapes in week B. No statistical procedures were applied to the results of these tests, since the intention in the main experiment was to measure progress not during the microteaching sessions but in terms of pre- and post-classroom performance. However, the assessments of the supervisory team indicated a generally high level of agreement, suggesting that a relatively short period of intensive training could produce reasonable reliability among the members of a small, committed group of raters.

(e) Modelling the skills

Borg et al (1970) present an interesting discussion (p. 67-68) of "Realism versus Pertinence" in the modelling of skills for microteaching. They began by recording classroom lessons taught by experienced teachers who had carefully worked through the analysis of the skill concerned, but it was found that even lengthy lessons of this kind yielded comparatively few explicit examples of the skill, embedded in a great deal of irrelevant material. It was therefore felt necessary to plan and even to script the material in a much more concentrated form, sacrificing naturalism to the needs of explicit exemplification. For some scripted models, actors were used in place of teachers. Experience at Jordanhill paralleled that of the Far West Laboratory. The department, aware of students' critical reactions to any hint of artificiality in teaching demonstrations, attached great importance to realism; but initial attempts by experienced lecturers to model the skills in the course of "normal" classroom lessons proved on the whole to be unsatisfactory. Careful analysis of the tapes revealed that each aspect of the skill in question had been used; but the pattern of moves and sequences/

sequences, of action and interaction, was so somplex that even patient editing could not make the examples of the skill stand out clearly from their context. Second attempts were more carefully structured and taken at a slower pace, though scripting was never adopted.

To achieve coherence in the presentation of the different skills aspects, passages of teaching were subsequently linked by studio commentary, recorded by another lecturer. Thus each completed model took the form of a unit integrating instruction with exemplification; a unit which could be used, if necessary, by students without support or interpretation by a tutor.

Research evidence (Orme, 1966; Young, 1967; Claus, 1969) pointed to the value of placing cues at appropriate points throughout the model tapes, to draw attention to the occurrence of the skills aspects. In the seven models prepared during 1971-72 (Clarity and Coherence, Question Techniques, Beginning the Sequence, Using Documents, Varying the Pace and Rhythm, Organization and Guiding Pupil Activity), the cueing took the form of verbal comments overlaying the soundtrack. Examples, taken from the model of Clarity and Coherence, were

Skill aspect	Comment
Clear voice	"Notice the teacher's clear, audible speech and his lively, confident style, capturing the pupils' attention."
Suitable vocabulary	"The teacher makes sure that the meaning of this technical term is understood."
Examples to clarify points	"The teacher clarifies the concept of a 'parliamentary session' by reference to an example within the pupils' experience."

This procedure provoked mixed reactions from groups of graduate students with whom the models were used for general instructional purposes, apart from the research study. A number of the students felt that the comments were pre-empting their own judgement of the quality

of/

of the lesson, thus reducing their role to that of passive acceptance of someone else's concept of effective teaching. Whether or not there would have been a similar reaction if the students had been viewing the models in the context of a microteaching session is open to question, since these tapes were not in the event used for microteaching purposes. Fresh models were prepared for the revised experimental design adopted in 1972-73 and the opportunity was taken to replace verbal cueing by captions bearing the name of each skills aspect, superimposed on the picture at appropriate points.

4. Method and procedure: the main experiment

(a) Amending the plan of research.

In setting out to examine the effects of microteaching on subsequent classroom performance, the research study initially aimed to gather evidence on a broad range of skills and to involve a high proportion of the available student population. Since there was only one researcher available for the study, this broad coverage could be achieved only through dependence on tutors' attainments of classroom teaching. The results of the June 1972 reliability trials, set out in Chapter 3, imposed a choice between postponing the main experiment for a year in order to carry out further training with the Appraisal Guide and revising the research design. A number of factors favoured the latter choice. It was clear that combining microteaching with classroom assessment of students in each half term period would make it very difficult to preserve confidentiality. In order to complete their schedule of school visits, tutors would need to use the mid-week and would thus almost certainly discover which students were still in school and which had been withdrawn to microteaching. Moreover, the indications of the pilot microteaching sessions were that students would have problems in assimilating and practising all aspects of the broad skills as defined in the Analysis. Finally, in the late summer of 1972, the History department heard that student numbers were likely to double in the following session. This would obviously impose a considerable extra burden on tutors, and the withdrawal of three lecturers from school visiting so that they could concentrate on microteaching appeared to be no longer feasible.

For all these reasons, it was decided to limit the scope of the study to investigating selected aspects of one basic skill and to restrict the microteaching practice to the second term, in order to allow time for/

for recording classroom lessons, taught by experimental and control groups in terms 1 and 3. The recordings would enable a category analysis of teacher-pupil interaction to be carried out, providing the data on which quantitative measurements of the use of the skills aspects would be based. At the same time, tutors could make their own assessments, using the revised form of the Appraisal Guide, of the students' use of the skills and their overall teaching competence.

The disadvantage of the new plan lay in the restriction it imposed on the range of skills to be investigated and - since only one researcher was available to carry out the classroom recordings - on the number of students involved. Nevertheless, the revision had a number of advantages. The prospect of achieving more reliable findings was important, not only for the sake of research principles but as a means of encouraging the adoption of innovative techniques within the History department. The principal lecturer pointed out that unreliable and inconclusive results would make it difficult for him to persuade the more sceptical among his colleagues to accept such innovations and he therefore welcomed the chance of obtaining more precise, even though more restricted evidence. Moreover, the new plan seemed likely to provide more valid evidence on the question of transfer of skills training from clinic to classroom. In the original design, classroom performance would have been measured only a week or two weeks after microteaching practice, whereas the new design incorporated assessment after the lapse of only a week. The new plan would also enable the researcher to combine and compare high and low inference assessments, along the lines recommended by Rosenshine and Furst (1971). The need to achieve greater reliability did not diminish the importance of retaining the lecturers' contribution. It was accepted that they alone were competent, as subject specialists, to assess appropriateness and discrimination in the use of the skills; and it was considered

essential/

essential to take appropriateness into account in assessing performance in the classroom situation. If the experience of microteaching were to engender in the student a dogged determination to use a particular technique more frequently, without regard to the varying needs and abilities of different classes, there would probably be little correlation between progress in microteaching sessions and progress in classroom effectiveness. Comparison between quantitative measurement and subjective assessment was therefore a basic element in the research design. Moreover, from an operational point of view, it was thought advantageous to maintain the lecturers' sense of involvement and commitment throughout the main stages of the research, in order to foster their interest in the new techniques of training which microteaching involved.

(b) The new research design

(i) Selection of skills

It was decided to concentrate the study on selected aspects of Question Technique. Questioning was chosen, partly because it lent itself to quantitative assessment, but principally because the History department regarded training in questioning skills as a priority in any future use which they might make of microteaching for remedial purposes. Poor questioning techniques were, in their view, a basic weakness among students.

The selected skills were grouped under two headings as follows:

Asking questions

1. a. Avoiding over-use of "one-word answer" type questions

b. Avoiding over-use of "yes/no" type questions

2. Varying the level and nature of questions, with particular attention to increasing the proportion of 'higher order' questions, demanding interpretation, judgment, etc.

Dealing with answers

3. a. Giving pupils time for thought in formulating asnwers

   b. Through prompting, encouraging adequately phrased and accurate responses

4. Making constructive use of initial answers by

   a. probing, for extension, clarification and justification of response

   b. re-directing initial answers for comment and discussion

This selection from the skills listed in the analysis of Question Technique was aimed at encouraging students to use questioning not merely as a means of promoting recall and obtaining information but as a stimulus to thought and discussion at a higher cognitive level. The justification for adopting this approach to questioning, outlined, for example, by Gall et al (1971, pp. 11-18), was accepted by the History department.

(ii) Introduction to microteaching procedures

In the introductory week of the autumn term, all honours and ordinary graduates in History (a total of 98 students) attended a lecture on microteaching. This lecture attempted to set the technique in the context of training as a whole; to discuss the identification and analysis of teaching skills; to give some account of the nature and development of microteaching; and to describe the 1972-73 research project and the role which students/

students were invited to play in it. The lecturer stressed that any performance measurements made in the course of the project would be used for research purposes only and would not form part of the formal assessment recorded on the teaching certificate; that invitations to take part in the project would be made on a random basis and would carry no implication of weakness in a particular skill; and that the maintenance of confidentiality was essential, to avoid any risk of bias in the evaluation of post-microteaching classroom performance. At the end of the lecture, all students were given a copy of the handout "Four Basic Questioning Skills" (see Appendix F). It was emphasized that this handout, which contained explanatory notes on the analysis and application of the selected skills, should be retained and consulted by students throughout the session. (Copies of the handout were also given to all History tutors.)

(iii) Selection of students

Because of the limitations of time and research assistance, the maximum number of students acceptable for the project was forty, divided into

Control group - 20 (10 in each half-term)

Experimental group - 20 (10 in each half-term)

It was clearly desirable that these students should share a common training experience throughout the year, with microteaching as the only independent variable. It will be recalled that graduates offering History as their sole "main subject" did all their training with the History department, while those combining History with a second subsidiary subject received their term 2 training/

training from their "Main Subject II" department. There were fewer than forty graduates in the former group, so it was decided to draw all members of the research population from the M.S.I/M.S.II graduates. The fact that these graduates would receive their term 2 training from different departments constituted an uncontrolled variable, but it was felt that this variable would be relatively unimportant. A considerable majority of the M.S.I/M.S.II students were offering Modern Studies, Economics or Geography as their subsidiary subject, and it was reasonable to suppose that the skills of questioning related to these subjects would contain many common elements. It was decided to exclude from the experiment the few graduates who were offering a modern language as a subsidiary subject, since it was argued that the approach to questioning skills implicit in the teaching of modern languages might constitute a significantly different experience from that of graduates who were doing all their training in the social subjects.

These decisions effectively reduced the available population to

First half-term - 31

Second half-term - 32

In each half-term, twenty students were randomly selected to take part in the experiment. These twenty students were subsequently matched in pairs, on the basis of their ability in the selected questioning skills, as assessed by tutors on their term 1 teaching practice visits. From each pair, one student was randomly selected to undergo microteaching training in term 2, the other member of the pair acting as the control.

It would have been simpler from a practical point of view to make a random division of each set of twenty students into experimental/

experimental and control groups, without attempting to match them for ability. However, it was considered important that the two groups should have an equal scatter of initial ability in the selected skills. At the end of the day, statistical adjustments could have been made to compensate for any significant differences in performance revealed at the pre-microteaching stage, but the results might still have been rendered invalid by the differing reactions of groups unequal in ability range to the actual experience of microteaching. For example, a group initially high in questioning ability might have made little progress as a result of microteaching, because in their case there was little room for improvement, while a control group of initially low ability might have shown considerable improvement throughout the year, simply as a result of gaining confidence through experience in the classroom. If the student population used in the study had been greater, the probability of obtaining groups approximately equal in ability through random selection would have been correspondingly higher; but with sets of only twenty students in each half-term, there was clearly a danger that random selection might produce two groups differing markedly in initial skills perfromance.

This line of thinking was close to that of Oppenheim (1970), who advocated that, in longitudinal studies, control and experimental groups should preferably be matched in pairs and that matching should be based on variables relevant to the dependent variable. It also found support in Kerlinger (1964), who accepted the advantages, in certain circumstances, of matching subjects on one or more characteristics, provided that members of each pair were assigned to one or other group at random and that a random decision was taken on which group was experimental and which was the/

the control. The new design satisfied these criteria and indeed conformed to Kerlinger's "classical" model of a pre- to post-test design:

Mr	Yb	X	Ya	(experimental)
	Yb	(~X)	Ya	(control)

where Mr indicates matching with some element of randomization

Yb indicates group at pre-test stage

Ya indicates group at post-test stage

X indicates manipulation of independent variable

(~X) indicates independent variable not manipulated

Since the research was concerned with the effect of microteaching on specific skills rather than teaching performance as a whole, it was decided that the base for matching should be performance on the skills in question. No other factor (type of degree, age, sex, etc.) should be taken into account, since it was not known which, if any, of these factors correlated with teaching ability. Given the small number of subjects, it seemed best to regard the students as a homogeneous group in all respects other than their ability in questioning skills.

Matching was based on tutors' assessments in term 1, partly because at that stage the reliability of the instrument for coding question categories had not been finally demonstrated and partly because the tutors' assessments provided evidence based on three sample lessons by each student. It was an unavoidable limitation of the design that the reliability of quantitative measurements was reduced by the fact that only one sample of each student's teaching could be recorded in each of terms 1 and 3. Moreover, tutors had shown themselves to be reasonably reliable in their assessments of specific aspects of questioning skills (see

Chapter/

Chapter 3, Table IV).

Matching was achieved by calculating average scores, based on the three tutor assessments, for each student on each of the seven selected aspects of question technique; by comparing each student with all the other students in turn; and for each comparison summing the squares of differences in scores between students on each of the skills. The ten closest comparisons (represented by the ten smallest sums of squares) were taken as the basis for matching the pairs of students.

(iv) The recording of classroom lessons

The terms 1 and 3 recordings of sample lessons by students in the control and experimental groups had to be fitted into each half-term period while avoiding, as far as possible, any clash with the tutors' schedule of visits. There was thus time to record only one lesson by each student in each term. Category analysis of performance would have been more reliable if it could have been based on a sample of, say, three lessons rather than one, but without assistance it was not possible to achieve this.

Classroom teaching involved a number of variables; for example, the social character of the school's catchment area, the age and ability range of the pupils, the subject-matter in question and the format of the lesson. In designing the research study, decisions had to be taken on the possibility and desirability of controlling these variables. The principal aim of the inquiry was to discover whether skills acquired through microteaching could be retained under normal classroom conditions; and there was

a/

a danger that, in manipulating those conditions in an effort to control or eliminate variables, the plan might create an artificial, self-conscious situation, thus reducing external validity. In the event, it was decided that all recorded lessons should be taught to first-year pupils, and schools were asked to provide mixed sex classes of average ability. It was the policy of the History department to change teaching practice schools for each student between terms 1 and 3, but an effort was made by the senior lecturer (in charge of school allocations) to keep the social character of the schools constant, between terms 1 and 3, for each student in the control and experimental groups. It would have been difficult to control the subject-matter of the lessons, since students on teaching practice were required to implement the scheme of work drawn up for the pupils, and this was not common to all schools. Principal teachers of History would probably have acceded to a request from the College to allow students to teach a "special" lesson on a designated topic for the purpose of the research project, but this would have created the flavour of the special occasion, which the researcher was anxious to avoid. In any case, the project insisted throughout on the importance of appropriateness in question technique and there was felt to be a close and logical link between appropriate use of questions and appropriate choice of subject-matter. Students were therefore left free to choose the topic for their recorded lessons, within the constraints imposed by the scheme of work operating in their teaching practice school. It is interesting to note that Acheson and Tucker (1971), in their research on factors affecting the acquisition of skill in asking higher cognitive questions, took the same decision for the same reason. They concluded "that allowing teachers to choose their own discussion/

discussion topics permits the important variables to operate in studies of this kind and that the disadvantages of assigning topics outweigh the advantages. The occurrence of higher cognitive questions may be more a function of the topic for discussion which was chosen than it is of the teacher's skill in using higher cognitive questions."

The students were, however, given guidance on format and asked to teach a 25-minute lesson, divided into

- 10 minutes of introduction and exposition
- 10 minutes of questioning and discussion
- 5 minutes of summing up

and to submit a lesson plan indicating these three main divisions. Subsequently, the ten minutes of questioning and discussion were identified and timed exactly from the recordings, and category analysis was applied only to this section of each lesson.

(v) Preserving confidentiality

By limiting microteaching practice to the second term, the new design made it easy to preserve the principle of confidentiality, since in that term all M.S.I/M.S.II students were working with other departments. Therefore, with the co-operation of those departments, the experimental group could be withdrawn from teaching practice to microteaching without the History lecturers being informed. Throughout the experiment, no History tutor making classroom assessments was aware of which students were taking part in the study. The tutors fully appreciated the importance of making blind assessments and accordingly made no attempt to question students about their participation in the project.

(vi)/

(vi) Organization of microteaching sessions

Microteaching practices were held in the mid-week of each half of term 2, with the following programme:

Day 1	Group A (5 students)	Skills for Asking Questions
Day 2	Group B (5 students)	Skills for Asking Questions
Day 3	Group A (5 students)	Skills for Dealing with Answers
Day 4	Group B (5 students)	Skills for Dealing with Answers

The timetable for each day's practice was

08.30 - 09.00	Set up recording and playback equipment
09.00 - 09.30	First pupil group collected from schools
09.00 - 09.30	Students view and discuss Model tape
09.30 - 10.15	3 students teach micro lessons
10.15 - 10.30	B R E A K
10.30 - 11.00	2 students teach micro lessons
11.00 - 11.15	B R E A K (First pupil group returned to schools)
11.15 - 12.55	Playbacks - tutor with students in group
12.55 - 14.15	L U N C H - students replan lessons - second pupil group collected from schools
14.15 - 15.00	3 students re-teach micro lessons
15.00 - 15.15	B R E A K
15.15 - 15.45	2 students re-teach micro lessons
15.45 - 16.00	B R E A K (Second pupil group returned to schools)
16.00 - 17.15	Playbacks - tutor with students in group

Pupils were drawn from two neighbouring schools. Headmasters were asked to supply first year pupils of average ability, with a mixture of boys and girls in each group of five. (Fresh pupils were used in each teaching session, making a total of forty pupils/

pupils per week.)

An appreciable number of research studies supported the value of cued perceptual modelling in microteaching work. Videotape recordings were therefore made, demonstrating the use of each of the skills selected for practice. Playback facilities were made available throughout each day of the microteaching sessions, so that, in addition to the initial morning viewing, students could study the skills for themselves while they were awaiting their turn to teach or re-teach. Accordingly, the recorded teaching extracts were linked by explanatory studio comment. (Scripts of the studio links are attached as Appendix G - Asking Questions and Appendix H - Dealing with Answers.) Experience had shown the need for careful structuring in exemplifying the use of skills, so the teaching sequences were recorded in the television studio, using small classes of ten to a dozen pupils. It was felt that the results were less "real" in atmosphere but more explicit in content than a classroom recording would have been. Since the first model (Asking Questions) dealt with the two avoidance behaviours of eliminating over-use of questions in a form that prompted single-word or yes/no answers, some use was made of negative illustrations, but each negative instance was followed up by an example of the teacher re-phrasing his questions to produce a fuller response.

For each session, students were reminded of the skills to be practised and given the following instructions:

> "By way of preparation for these microteaching sessions, you should study the "Specific Analysis" section of your notes on Four Basic Questioning Skills. Then choose a history topic suitable for first year pupils and plan an eight to ten-minute treatment which allows you to practise the specific skills of questioning. Obviously, any teaching sequence, if it is to be coherent, will include elements of other skills, such as/

> as introduction, exposition and summing-up; but keep these aspects of your teaching as brief as possible and concentrate on your question technique."

No specific lesson topics were set, for the reasons given in sub-section (iv) above.

In view of the fact that most of the research studies concerned with recorded feedback (e.g. Klingstedt, 1970; Acheson and Tucker, 1971; Bortz, 1971; Gall, 1971) had found no significant difference between audio and video recording, audiotape was used for recording the student lessons. The equipment was easy to rig, simple to operate and relatively unobtrusive.

One History tutor, selected from the microteaching team trained in the preparatory year, acted as supervisor throughout the sessions. This had the practical advantage of freeing the maximum number of lecturers for school visiting and also eliminated the variable associated with different styles of supervision. Research evidence (Johnson and Knaupp, 1970; McIntyre, 1971) indicated that students welcomed the presence of a supervisor at the feedback stage and that they gained more from indirect than direct supervisory styles (Blumberg, 1965, 1968; Johnson, 1967; Johnson and Knaupp, 1970). A briefing letter to the supervising tutor stressed the following points:

- the importance of basing comments upon the definition and analysis of skills which had been given to students in the handout
- the wisdom of limiting discussion immediately after viewing the models to points of clarification and reinforcement. (Too much discussion at this stage might confuse students about to teach their micro-lessons)
- the need to encourage each student, at the review of the teaching stage, to evaluate his own performance on the evidence of the tape replay. The tutor's role at this juncture should be supportive rather than directive
- the/

- the value of ensuring, at the reteach stage, that each student fully understood the nature of the skills he had been practising and fully grasped the major points of strength or weakness in his performance.

It was decided (on the evidence of student reaction to other reported experiments - e.g. Owens and Hatton, 1970; Turney, 1970; Young D A, 1970) to conduct the playbacks as group sessions. Students were given an Evaluation Form (see Appendix I) to complete for each playback, as a means of directing their attention to the various aspects of the skills being practised.

Since there was time for only one sequence of training on each of the skills aspects, it was thought best to operate the complete teach - reteach cycle, although some studies (Orme, 1966; Skailand, 1972) have cast some doubt on the value of the reteach element.

(vii) The measurement of results

To fulfil the aims of the research study, it was necessary to

- compare the term 3 performance of the experimental and control groups in the selected questioning skills
- compare the quantitative measurement of those skills with the assessments made by tutors
- relate the experimental group's performance in the specific skills to their overall lesson performance as assessed by their tutors
- analyse the reactions of students and tutors to the experiment.

In determining procedures for the analysis of results, the prior need was to ensure, as far as possible, that any measurements used had a proven basis of reliability and that good statistics were not being wasted on poor data. The problem of developing reliable/

reliable measuring instruments and of training raters in their use extended beyond the tutors' Appraisal Guide to the coding instrument needed to categorize questioning behaviour, and the preparation and testing of the coding system will be described in chapter 5 below. Another source of unreliability, perhaps characteristic of small-scale research studies, lay in the restricted population (forty students divided into groups of ten) and the limited number of samples on which measurements could be based. Thus some of the skills on which evidence was sought were practised so rarely that it would have been difficult to base statistical procedures upon the data available and dangerous to place too much reliance upon any results obtained. Moreover, the variables associated with day to day classroom teaching - size and ability of pupil groups, interest of subject-matter and so on - introduced en element of random error which could have been controlled only by imposing upon the design a degree of artificial structure which would have reduced its external validity.

These limiting factors, which arose partly from the operational nature of the research, placed some restrictions on the development and refinement of statistical procedures. It would have been interesting, for example, to analyse the occurrence of Probes separately for each of the three broad cognitive categories (Higher, Middle and Lower Order), but the totals for the separate categories were so small that it would have been unwise to attach significance to any resultant comparisons. Similarly, the skill of Redirection of responses was practised so rarely overall that it was not possible to make any reliable comparison of mean performance. It appeared likely that the only behaviours which could meaningfully be compared were those related/

related to broad categories which would provide sufficient data for reliable analysis.

Emphasis was also placed on measuring classroom performance rather than progress in the course of the microteaching cycle. The evidence of previous research studies, that students given microteaching practice could significantly increase their use of specific skills within the microteaching context, was accepted by college tutors. Their interest lay entirely in discovering whether a limited experience of microteaching could significantly affect subsequent classroom teaching behaviour.

With these considerations in mind, the following measurements were adopted as part of the research design:

(1) Quantitative measurements

To determine the significance of differences between groups, in respect of mean scores obtained for each of the selected skills, t-tests were carried out. The test was considered appropriate in that the primary concern was with comparison of means; and it also had the advantage of robustness with respect to normality of population distribution. Significance was examined in relation to

Term 3 performance	—— Term 1 performance	Experimental (1-tailed test) Control (2-tailed test)
Experimental group	— Control group	Term 1 (2-tailed test) Term 3 (1-tailed test)

and particular importance was attached to the comparison between/

between experimental and control groups in term 3. It was considered justifiable to apply a 1-tailed test to the term 3 - term 1 experimental group comparison and to the experimental - control comparison in term 3, since it seemed reasonable to hypothesize, on the strength of previous research, that the microteaching variable would have a positive effect on performance.

It was also decided to apply t-tests of significance to differences in performance gains, by examining

Experimental group gains —— Control group gains
Term 1 → Term 3 Term 1 → Term 3

It was appreciated that a comparison of differences in performance gains would be a less reliable measurement than direct comparisons between performance means, but it was thought worthwhile to include this test, as a means of checking whether, for a specific skill, the difference between experimental and control groups in term 3 was significantly greater than the difference in term 1.

In order to make a graphical comparison of group performance on each of the skill categories, the raw scores were converted to standard scores, and these were used to plot performances about a zero mean.

An analysis of variance was carried out on all term 3 scores, to examine the interactive effects of the different factors which could have affected performance, namely

- differing treatments
- membership of the first half term or second half term teaching practice group
- initial ability on the selected skills.

Three/

Three ratios were tested for significance:

$$\frac{a}{b} \quad \frac{\text{Variance due to first half/second half term effect}}{\text{Variance between pairs within half terms}}$$

Within pairs

$$\frac{c}{e} \quad \frac{\text{Variance due to differing treatments}}{\text{Variance due to treatment x pairs within half terms}}$$

$$\frac{d}{e} \quad \frac{\text{Variance due to treatment x half terms}}{\text{Variance due to treatment x pairs within half terms}}$$

It was also decided, for each half term section, to express the differences between control and experimental groups in terms of the number of skills on which each group showed

- superiority in terms 1 and 3
- superior term 1 to term 3 gains.

The probability factor related to each of these comparisons of superiority was then determined and examined for significance.

It was hypothesized that increased efficiency in the selected questioning skills should result in a reduction of the percentage of teacher talk. (Borg, 1970, had reported a highly significant reduction of this percentage in his analysis of the results obtained from the main fiels test of Minicourse 1: Effective questioning at elementary level.) Students were given no direct instruction to reduce their amount of talking in the recorded classroom discussions, since this might have resulted in the employment of devices - pausing for excessive periods, for example, or addressing all questions to a few talkative children - which would have turned the lesson into an artificial exercise. The reliability of the procedure for measuring teacher talk had already been established at the University of Stirling by

two/

two other researchers who had achieved correlations averaging .97. These researchers reported that the problems involved in achieving reliability were mainly technical, concerned with the accuracy of stop clocks and the consistency of tape recorder running speeds. The only other problems related to decisions about pauses and overlapping speech. After discussion with the researchers, a set of ground rules for assessing the proportion of teacher talk in recorded classroom lessons was drawn up for the Jordanhill project (see Appendix O, Section 2). The assessor tested these ground rules on four 5-minute extracts randomly selected from a series of recorded student discussion lessons. Two measurements, with an interval of several weeks between them, were made by the assessor on each extract. The results were

TABLE IX

Reliability trial of teacher talk measurement

	Duration of teacher talk	
Lesson	1st assessment	2nd assessment
1	4'13	4'13
2	3'84	3'81
3	3'40	3'40
4	4'19	4'17

On the basis of these results, it was considered legitimate for the same assessor to claim reliability for his measurements of teacher talk in the term 1 and term 3 recorded classroom lessons, and to base comparisons between experimental and control groups on these measurements.

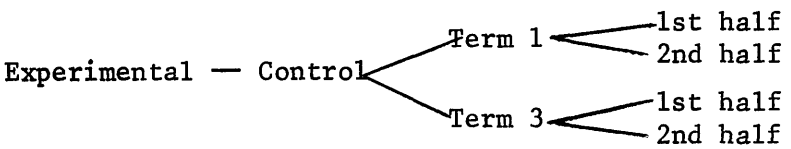
In/

In each half of term 1, the experimental and control groups were ranked for ability in the selected questioning skills, on the basis on the tutors' assessments. It was decided to correlate these assessments with the quantitative measures of skills performance derived from the term 1 recorded lessons. It was considered that the level of these correlations would give some indication of the degree of consistency with which students employed the skills, and possibly provide a pointer as to whether the concept of appropriateness in the use of skills constituted a significant variable between tutors' assessments and quantitative measures. It was realised that the level of correlation would not in itself either prove or disprove the reliability of the matching procedure, since it was likely that the uncontrolled variables in the day to day teaching situation would have an appreciable effect on a student's use of skills from one lesson to another.

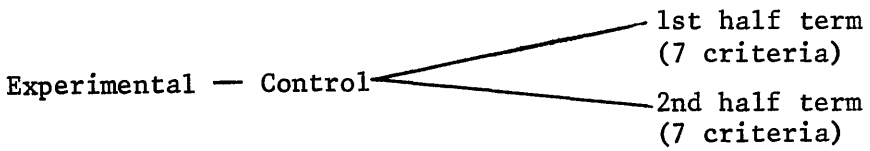
(2) Tutors' assessments

Using the data available from tutors' Appraisal Guide assessments of classroom lessons, t-tests were applied to discover whether any significant differences emerged between experimental and control group performances:

Ratings of the lesson overall



Term 3 ratings of separate skills



The tests applied to the overall lesson ratings were 2-tailed, since/

since there was no available evidence to support a positive hypothesis that microteaching training in specific skills would raise the quality of teaching performance as a whole. The tests in respect of specific skills were 1-tailed, on the assumption that microteaching work on those skills would have a beneficial effect on their subsequent classroom application.

The History department agreed that, if their assessments provided a different indication of student progress from that of the quantitative measurements, pairs of tutors would make two independent assessments of questioning skills performance in each of the term 3 recorded lessons. This would provide a more reliable basis for comparison, since both high and low inference assessments would relate to the same lessons. These tutors' assessments of recorded lessons would carry one of three implications:

- if there was a large measure of disagreement within pairs of tutors, their ratings would be demonstrably unreliable and there would be no point in pursuing statistical comparisons of high and low inference approaches;

- agreement within tutor pairs but disagreement with the quantitative measures would indicate that tutors were taking account of factors (such as apporpriateness and use of opportunities) which could not be measured quantitatively, and that these factors constituted a significant variable in assessment procedures;

- agreement both among tutors and between tutors and the quantitative assessor would suggest that students were capable of using specific skills in the classroom when they felt they were required to do so (e.g. on the occasion of a specially recorded lesson), but were not necessarily incorporating them into their teaching technique on other occasions.

Assuming that inter-tutor agreement was achieved, the following product-moment correlations would be elicited for/

for each of the skills categories:

Quantitative measure of recorded lesson skill	⟷	Tutor A rating of same skill	
Quantitative measure of recorded lesson skill	⟷	Tutor B rating of same skill	
Quantitative measure of recorded lesson skill	⟷	Tutor's rating of classroom lessons by same student	
Tutor A rating of recorded lesson skill	⟷	Tutor B rating of recorded lesson skill	} for each pair of tutor ratings
Tutor A rating of recorded lesson skill	⟷	Tutor's classroom rating of same skill by same student	
Tutor B rating of recorded lesson skill	⟷	Tutor's classroom rating of same skill by same student	

These correlations would be based on the whole population of students in term 3 (i.e. 40 students), since the concern was with comparison of measurement procedures, not with differences of performance.

(3) Analysis of student and tutor reactions

Apart from notes taken at informal discussions with students and tutors at seminars and staff meetings, analysis of reactions to the microteaching experiment was based on three questionnaires

- for completion by the experimental group of students immediately after their microteaching sessions in term 2;

- for/

- for completion by the same students immediately after their final teaching practice in term 3;
- for completion by tutors at the end of the session.

Copies of the questionnaires are attached as Appendices J, K and L. The main purpose of the term 2 student questionnaire was to gather reactions to the processes of microteaching while the experience was still fresh in the mind; whereas the term 3 questionnaire called for a retrospective assessment of the value of microteaching seen in the light of subsequent teaching practice. The tutors' questionnaire was designed in three sections aimed at sounding opinion of the conceptual basis of microteaching, gathering reactions to the use of the Appraisal Guide and inviting suggestions for future uses of microteaching.

In planning the form of the questionnaires, the researcher had to balance the known reluctance of both students and lecturers to spending much time on completing this kind of document against the danger of phrasing questions in a simplistic or ambiguous form. The guide-lines adopted were

- wherever possible, the questions were presented as multiple choice items, and response was limited to a simple tick or ticks
- space for additional comment was left at intervals throughout the questionnaires, to balance the closed form of the main questions
- many of the questions were preceded by structuring comments to recall the circumstances or concepts to which the questions related
- attitude statements were included, built up from the kind of reaction statements most frequently noted at informal discussions
- care was taken to avoid leading or loaded questions
- questions implying an attitude scale were set down/

down in random order, avoiding a fixed progression from "favourable" to "unfavourable".

In view of the small student and lecturer population, no attempt was made to design the questionnaires in a form adapted to sophisticated statistical procedures.

(c) Operation of the experiment

The results of the final lecturers' reliability trials in September 1972 underlined the difficulty of achieving high inter-rater reliability on the basis of high inference evaluation. The group had worked very hard to agree on norms of performance in the teaching skills; but it appeared that, even when agreement was achieved in relation to a particular lesson, there was a problem in transferring this agreement to other lessons taught by other students. Change in personality, teaching style, lesson objectives and subject-matter, class ability and so on introduced so many variables that generalisation became almost impossible. One could certainly hope that out of discussion and analysis certain broad lines of agreement would emerge; for example, that too much teacher talk, too little mental involvement by the pupils, too little variation of stimulus would be characterised as weaknesses in <u>any</u> teaching situation. The problem, in the context of a specific lesson, was to define the levels of behaviour which constituted "too much" or "too little".

Some guidance in this task was offered in the form of "Hints for completing the Appraisal Guide", circulated to lecturers at the beginning of term 1 and reproduced as Appendix M(a). In summary, these Hints sought to establish the concept of weak → average → very good performance,/

performance, as a basis for assessing all observed lessons; to urge that this concept should be applied as objectively as possible, without making allowances for the experience or personal problems of any particular student; and that "appropriateness" should be taken into account in evaluating the use of skills. Characteristics of "below average/average/above average" performance were set out as a guide, in relation to each of the selected behaviours.

These hints were supplemented towards the end of the term by a note on avoidance behaviours (see Appendix M(b)) and by a letter of reminder (Appendix M(c)) sent out before the beginning of term 3.

It was hoped that, with the help of this briefing, the level of reliability would remain at least as high as that achieved in September 1972. It would have been helpful to have held further group replay and discussion sessions as a reinforcement in term 2, but no opportunity was found for arranging them. Throughout the session, Appraisal Guides were completed and promptly handed in by all lecturers in respect of each school visit. Some lecturers added a good many notes, others restricted themselves to the assessments alone.

Briefing letters were also sent in terms 1 and 3 to all students selected to take part in the project. These letters, which included details of lesson duration, planning and structure, and which stressed the fact that evaluation of recorded teaching would form no part of any "official" assessment, are reproduced as Appendices N(a) and (b).

All but one of the graduates co-operated willingly in the project. The single objector was concerned that the experiment should "place on record an example of how a lesson should not be taught".

All the term 1 and 3 classroom lessons were recorded by the researcher on audiotape. Audio recording seemed justified for the reasons outlined

in section (b)(vi) of this chapter. Since the researcher had to record an average of three lessons a day in different schools, it was stipulated that the equipment used should be

- light enough for one person to carry in or out of a school in one journey
- capable of being rigged or de-rigged in 3-5 minutes
- sufficiently straightforward for one person to operate.

It was necessary, to fulfil the requirements of the coding instrument, that, in addition to the teacher's voice, at least 80 per cent of pupil responses should be intelligible. This was achieved by the use of a lightweight boom microphone, held and pointed by the recordist.

The class microphone was activated by a switch at the base of the boom, which (to allow for overlapping speech) attenuated but did not cut out the lanyard microphone worn by the teacher. It was recognized that even audio recording would introduce an element of artificiality into the teaching situation but this was held to be a controlled variable, common to all the lessons. All schools gave ready assistance in planning these recordings and a number of principal history teachers showed considerable interest in the project.

Co-operation came equally readily from the heads of the College departments who were asked for permission to release experimental group students to microteaching for two days in the second term. Several heads stated that the questioning skills to be practised were equally relevant to the teaching of their own subject.

These microteaching sessions ran smoothly, without any great technical or logistical problems. Students were fully briefed and reminded of the need for confidentiality. It was of interest that one student who, assuming that videotape was to be the recording medium, requested to withdraw on the grounds of acute anxiety, was quite reassured when she learned/

learned that audio recording was to be used. One variable which had not been foreseen was the different reaction of morning and afternoon groups of pupils. Without exception, the morning pupils appeared more alert and readier to respond than the afternoon pupils, and a number of students commented that this made the reteach lesson harder work than the initial lesson. If the teach and reteach lessons had been submitted to category analysis, this variable might have produced a misleading impression of the value of the reteach stage. The students' own scoring of peer lessons showed that, in their judgments, there was a slight gain in effective use of the skills in the reteach lesson, but no reliance can be placed on these assessments.

Reactions of students and lecturers to the experiment as a whole are discussed in Chapter 6.

5. Preparing and testing the Lesson Coding Instrument

In order to measure the students' use of the selected questioning skills in quantitative terms, it was necessary to devise a coding instrument and demonstrate its reliability. This involved providing an exact definition of each behaviour and precise instructions for recording its occurrence. These definitions and instructions were incorporated in "Ground Rules for Analysis and Coding of Student Lesson Excerpts" (see Appendix O). The coding tasks related to each of the skills were as follows:

(i) Avoiding over-use of "one-word answer" type questions

(ii) Avoiding over-use of "yes/no" type questions.

Question types were defined, and their occurrence in each 10-minute lesson excerpt was enumerated.

(iii) Varying the levels and nature of questions, with optimum use of higher order questions.

A taxonomy of question levels was established and each question was recorded in its appropriate category.

(iv) Giving pupils time to think out adequate answers.

This was a difficult skill to assess in quantitative terms, since it could reasonably be argued that the time needed for adequate answers varied with each question and each pupil, and was in any case dependent upon one's conception of adequacy. Measurement was therefore made by deciding on the minimum pause following a teacher question (and before the teacher resumed speaking) which would conceivably allow pupils to respond; and by coding all instances where such a pause did not occur.

It was also possible to obtain an indication, though

not/

not a direct proof, of this skill from measuring the percentage of teacher talk; a measurement which reflected the percentage of the excerpt occupied by pausing to allow pupils time to respond and by the responses themselves.

(v) Prompting, to encourage accurate, well-formulated answers.

(vi) Probing, to extend, clarify or justify initial responses.

(vii) Re-directing initial answers for comment and discussion.

Prompting, probing and re-directing were defined and the frequency of use was tallied for each excerpt.

In addition, questions to which no answer was obtained, even after a pause for response, were separately coded, in order to ascertain whether an increase in the proportion of higher order questions was associated with a reduction in pupil responses.

Direct student-to-student comparison, based on the frequency of each coded behaviour, would have been misleading, since the microteaching training programme did not aim to increase the total number of questions asked in a given period. It was concerned to bring about proportionate increases in certain behaviours and proportionate decreases in others. Accordingly, it was decided to express the data derived from the coding in terms of the following proportions:

(i) proportions excluding questions where no opportunity was given for response

$$\frac{\text{number of higher order teacher questions}}{\text{total number of teacher questions}} \qquad \frac{HOQ}{Q}$$

$$\frac{\text{number of lower order teacher questions}}{\text{total number of teacher questions}} \qquad \frac{LOQ}{Q}$$

$$\frac{\text{number of teacher questions in a form demanding no more than a single-word answer}}{\text{total number of teacher questions}} \qquad \frac{S}{Q}$$

number/

number of teacher questions in a form demanding no more than a yes/no response / total number of teacher questions	$\frac{y}{Q}$
number of teacher prompts following questions which attracted no response / total number of teacher questions which attracted no response	$\frac{XP}{X}$
number of teacher probing questions / total number of teacher questions	$\frac{Pr}{Q}$
number of teacher re-directions / total number of teacher questions	$\frac{R}{Q}$
number of teacher questions attracting no response / total number of teacher questions	$\frac{X}{Q}$
number of higher order teacher questions attracting no response / total number of teacher higher order questions	$\frac{HOX}{HOQ}$

(ii) proportion including questions where no opportunity was given for response

number of teacher questions where no opportunity was given for response (n) / total number of teacher questions, including n questions	$\frac{n}{Q \text{ incl. } n}$

"No opportunity for response" questions were excluded from all but one of the proportions, since the primary concern was to compare attainments in desired behaviours. If n questions had been included in determining the proportions of higher order questions, prompts, probes and re-directions, certain students who were often guilty of giving no opportunity for response might have appeared on paper to do well, although their teaching behaviour would in fact have been unlikely to lead to effective pupil response and discussion.

In working out the proportion for prompts, the total number of teacher questions which attracted no response was taken as the denominator, to avoid misleadingly low proportions in those cases where nearly all the teacher's questions were answered without the aid of prompting. In considering/

considering "no response" questions, a separate proportion was determined in respect of higher order questions, to provide evidence relating to a possible link between higher order questioning and failure to respond.

Since the chief aim of the training programme was to encourage students to use questions as a stimulus to thought and discussion at a higher cognitive level, the main task in developing the coding instrument was to categorize questions in terms of the level of cognitive activity which the teacher apparently intended to provoke in his pupils. The decision whether to adopt an existing taxonomy (see, for example, those listed in Chapter 2, section (a)) or to devise a new set of categories was influenced by a number of factors, namely

- the categories should bear some relation to the descriptive framework used by History department tutors in discussing questioning skills. It would be confusing for students if the microteaching programme, set in the second term, used a set of categories completely different from those employed in earlier lectures and seminars. In fact, the History department, in its references to question technique, used terms based on Bloom's cognitive categories.

- the categories had to lend themselves to grouping under broad headings, since neither tutors in their appraisal of student lessons nor students themselves in their training programme could be expected to assimilate and apply a complex taxonomy.

- the categories had nevertheless to be sufficiently precise to yield reliable measurements and detailed enough to portray behavioural differences with some sensitivity. It would/

would have been easy, for example, to group all questions under broad headings such as "factual" and "others", and raters employing such a crude distinction could probably have achieved a high degree of reliability; but this kind of simplification would have provided little information about the behaviours which the project wished to encourage.

- the categories should preferably be in line with those already in use by other researchers at the University of Stirling. Employing standardized coding procedures whose reliability had already been demonstrated would simplify the task of training raters and legitimately reduce the need for extended reliability trials. The lesson sampling instruments on questioning skills, produced at Stirling by White (1972) and associates, employed some of Bloom's categories, and the development work based on Minicourse 9 (Higher Cognitive Questioning) also employed a taxonomy almost identical to that devised by Bloom (1956).

These factors, taken together, led to the adoption of the following categories in the Jordanhill coding instrument:

A. Lower Order questions

factual information

straightforward recognition

simple definitions

establishing pupils' range of knowledge, understanding, etc.

unsupported opinion and unconsidered reaction

random guesswork

B./

B. Middle Order questions

providing a description or account

translating from symbolic to verbal form

straightforward comparisons

providing explanations in own words, of material being studied

making summaries

predictions based directly on apparent trends (i.e. extrapolation)

application of acquired information, principles, rules

C. Higher Order questions

identifying motives or causes

drawing conclusions

supporting generalisations

distinguishing objective statement from subjective opinion

identifying elements of agreement and disagreement in sources of evidence

making inferences, predictions, hypotheses

solving problems

producing original communications

opinions based on reasoned judgment

judging validity of ideas

evaluating merits of solutions to problems

arguing conclusions

D. Procedural and rhetorical questions

These categories, defined in detail in the Ground Rules, were referred to in the Notes for Students (see Appendix F) under the simple headings of Recall, Comprehension and Thought. It will be seen that the categories correspond/

correspond as follows to Bloom's taxonomy:

Lower Order	Knowledge
Middle Order	Comprehension
	Application
Higher Order	Analysis
	Synthesis
	Evaluation

Bloom's taxonomy was based on the assumption that "essentially the same classes of behaviour may be observed in the usual range of subject-matter content, at different levels of education ... and in different schools". An analysis of recorded student history lessons suggested that the taxonomy was indeed relevant to questioning in history teaching to secondary pupils, though certain problems, discussed below, became increasingly obvious. It was, however, felt necessary to develop categories relating specifically to the aims and content of historical inquiry, and here the objectives listed by Coltham and Fines (1971) provided a useful guide. Examples, at the higher order level, were

- distinguish objective statements from subjective opinions in secondary source material

- identify elements of agreement or disagreement between two pieces of historical evidence or secondary source material.

Moreover, all examples of questions in each category had a historical content. Wherever possible, examples were drawn from the stock of recorded lessons available to the researcher.

Many of the problems associated with the attempt to analyse teacher/pupil interaction/

interaction and to place questions in cognitive categories are reflected in section 5 of the Ground Rules, "General Coding Rules relating to Questions". The well known difficulty of capturing the complex, volatile pattern of classroom interaction within the framework of a category system was encountered early on in the development of the Ground Rules. No sooner had one problematical dragon been trapped by a hastily improvised sub-paragraph than a couple more sprang up unpredictably in its place. The temptation was to go on and on adding to and refining the rules, to cover all possible eventualities, until the system became so complex that it was extremely difficult for a coder to find his way through the maze of sub-rules and cross-references. The alternative strategy - to keep the rules clean and simple and to allow each team of raters to work out its own formula for applying them, through a process of repeated trial and discussion - was thought likely to be more effective in circumstances where one could rely on building up a close-knit group of coders and retaining them as a team throughout the whole period of a research study. This capacity of a team to develop an almost intuitive sense of common direction in decision making was noted by Bellack (1966), whose researchers, working together on various stages of content analysis, "continuously interacted, discussed problems of coding, and probably developed, over time, a shared perspective that in part accounted for the high agreement in the final coding". However, research within colleges of education can rarely afford the services of a team; and it was therefore felt necessary to write the Ground Rules for the Jordanhill project in some detail, so that they could be used reliably by other researchers wishing to replicate the study.

Early attempts at categorisation also made it plain that questions in the context of history teaching did not always fit neatly into distinct cognitive levels. On the contrary, questions appeared to lie along a continuum in terms of cognitive activity, and it was particularly difficult to judge the point at which "exploratory" questions, prefaced by phrases such/

such as "What would ....?" "When would ....?" etc, moved from the level of application to the level of higher order analysis/synthesis. Sanders (1966), working with Bloom's taxonomy, described this difficulty well:

> "The relationships between the categories of questions are similar to that between colors on a spectrum. There, the colors of red, orange, yellow, green, blue, indigo and violet are plainly visible. Between each color, however, is an area that is neither one nor the other, but a part of both. The same seems true of the categories in the taxonomy. The important point for teachers to remember is that difficulty in classifying any question is no detraction from the quality of the question."

Bloom himself acknowledged that "we have not succeeded in finding a method of classification which would permit complete and sharp distinctions among behaviours." Clearly, these grey areas between categories were a potential source of disagreement among coders. It was possible to reduce the grey areas by increasing the arbitrariness of the coding system, by declaring that a question in a certain form should always be classified in the same way. For example, Gall et al (1971), in their Summary Chart of Question Types, emphasize question stems as a guide to classification. But analysis of lessons provided many examples to show how misleading arbitrary guides to classification could be. Content and context, as well as form, had to be taken into account if a realistic assessment of cognitive level was to be made. It seemed likely that, as one moved away from sensitive but tentative assessments in the direction of arbitrary decisions, the gain in reliability would be offset by a loss in validity. Bellack, again, appreciated the problem in his work on classroom verbal interaction:

> "In many instances, solution of a problem (of categorization) depended upon a compromise between representing precisely the unique qualities of particular statements and establishing more abstract categories that lost some of the information conveyed by specific statements but also provided a reliable basis for classifying the meaning expressed. This is a problem encountered in almost all content analysis procedures, and these efforts represent a series of compromises which seemed to capture best the meanings expressed in classroom interactions."

The policy adopted in framing the Jordanhill Ground Rules was to reduce the/

the grey areas in the cognitive continuum as much as possible, by paying particular attention, through definitions and examples, to the type of questions that lay on the borderline between recognizable categories. At the risk of complicating the Rules to the point where the coder found them difficult to handle, arbitrary procedures for classification were kept to a minimum.

Another problem recognized by a number of researchers in this field was whether to classify questions in terms of the teacher's intention in asking them or to code them on the basis of the work actually done by the pupils who responded. Bloom had noted the problem in the introduction to his Taxonomy:

> "One of the major problems in the classification of test items which this study revealed is that it is necessary in all cases to know or assume the nature of the examinees' prior educational experiences. Thus, a test problem could require a very complex type of problem-solving behavior if it is a new situation, while it may require little more than a simple kind of recall if the individual has had previous learning experiences in which this very problem was analyzed and discussed. This suggests that, in general, test material can be satisfactorily classified by means of the taxonomy only when the context in which the test problems were used is known or assumed."

If "question" is substituted for "test problem", these remarks apply very well to the Jordanhill study. The difficulty was that the coder of the recorded classroom lesson had no sure means of judging the pupils' previous learning experiences or, indeed, the teacher's intentions in framing a question; since those intentions would be influenced by knowledge or assumptions (for example, about the content of previous lessons) available to the teacher but not available to the coder. In order to achieve reliability, it was essential that admissible evidence on factors influencing the teacher's intention should be restricted to data equally available to all coders. In other words (to quote the Ground Rules), "the criterion should be the apparent intention of the teacher in asking the question, as discernible from the form of words in which the question

is/

is expressed and the development of the lesson prior to the question, insofar as this provides evidence of the pupils' level of information, degree of understanding, etc." This was to apply an arbitrary procedure, but it was felt that in this instance arbitrariness was unavoidable. However, to avoid the risk of fostering judgments unrelated to reality, the Ground Rules went on to remind coders that content and context, as well as form, must be taken into account. A question, for instance, asking

"What would happen to your bicycle tyres if you rode over a lot of broken glass?"

would appear to require no more than the direct application of knowledge previously acquired. It could not reasonably be compared, in terms of cognitive activity demanded, to a question identical in form but different in content, for example

"What would happen to the British economy if we were to withdraw from the Common Market at this stage?"

Clearly, it would have been easier to achieve reliability if the Ground Rules had stated categorically that all questions in the form of "What would happen if ...?" should be given the same classification. The researcher deliberately chose to risk a lower degree of reliability for the sake of remaining more sensitive to the reality of cognitive activity demanded in each instance.

It was necessary, before formal trials with the Ground Rules commenced, to set standards for reliability and to decide on a formula for measuring agreement among coders. Training sessions in the use of the Rules were organised for three coders; the researcher and two members of the Education department at the University of Stirling, one of whom was engaged on a study related to higher cognitive questioning and had devised ground rules similar in many respects to the Jordanhill rules. It was decided that reliability trials should be based on the coding of five-

minute excerpts (concentrating on questioning and discussion) of at least ten student recorded lessons. Analysis should be made directly from the audio tapes rather than from transcripts, since lack of time and clerical support would subsequently preclude the use of transcripts in analysing the control and experimental group classroom lessons.

The choice of formula to measure agreement among coders depended partly on what aspects of agreement the research wished to demonstrate; partly on the structuring of the data; and partly on the degree of rigour which seemed appropriate. Examination of formulae used in other reliability tests suggested that it was difficult to devise a statistical procedure which accurately and sensitively reflected not only the occurrence but the nature of disagreements. Indeed, any such procedure would depend on prior decisions about relativity in degrees of disagreement, and there appeared to be no objective evidence on which decisions of this kind could be based. Thus, with three coders, a, b and c, working within the three broad question categories established by the Jordanhill Ground Rules, the following range of agreement/disagreement might appear in a sequence of ten questions classified as Higher Order by coder a:

	1	2	3	4	5	6	7	8	9	10
Higher Order	abc	a	a	ab	ab	a	ab	a	a	a
Middle Order		bc		c		b		b or c		
Lower Order			bc		c	c			b or c	

Only one question attracts overall agreement. Of the remaining nine, it could be reasonably argued that the degree of disagreement for 3 was greater than that for 2 or 4. But does 3 represent a greater or smaller degree of disagreement than 6? Does 4 represent more or less agreement than 7, where one coder has failed to record the question? And how does one relate 10 to 8 or 9? Patterns of this kind suggest that any formula

selected/

selected to measure agreement must be to a certain extent arbitrary and approximate.

This type of approximation to a "degree of agreement" could be obtained from the formula used by White (1972). He expressed the agreement among three coders, a, b and c, in terms of the percentage

$$100 \times \frac{3\,(\Sigma abc) + (\Sigma ab + \Sigma ac + \Sigma bc)}{\Sigma a + \Sigma b + \Sigma c} \quad \frac{\text{agreements}}{\text{total entries}}$$

If, however, the aim was simply to reflect the degree of overall agreement among three coders, one could use the much more rigorous formula

$$100 \times \frac{\Sigma abc}{\Sigma abc + (\Sigma ab + \Sigma ac + \Sigma bc) + (\Sigma a + \Sigma b + \Sigma c)} \quad \frac{\text{actual overall agreements}}{\text{possible overall agreements}}$$

Alternatively, one could look separately at a-b, a-c and b-c agreements, and here again one could choose between a "lenient" formula (as used by White in the later stage of his research)

$$100 \times \frac{2\,(\text{number of agreements})}{\text{total number of entries by both coders}}$$

and a more rigorous formula

$$100 \times \frac{\text{number of actual agreements}}{\text{number of possible agreements}}$$

For the Jordanhill project, it was decided to measure agreement between each pair of raters in turn; partly to discover whether the similar work being done by one of the Stirling coders would influence the level of agreement and partly because any measurement which combined the entries from all three coders raised problems of "degrees of disagreement" as described above; problems which the researcher was unable to resolve to his satisfaction.

A ratio of actual to possible agreements was preferred as being more logical even though it was certainly more rigorous, than a ratio of agreements to total entries. The difference becomes apparent if one considers a sequence of questions recorded by coders a and b.

Higher/

Higher Order		a	ab	b		ab	a	
Others	ab	b		a	ab		b	ab

In this sequence, the formula used by White would give a ratio of 4 (2 x number of agreements) to 7 (total entries in Higher Order category) for Higher Order agreements. The actual to possible agreements ratio would be 2 to 5. The former measure implies that

> "coders a and b between them recorded 7 Higher Order questions. 4 out of the 7 entries were in agreement."

This would appear to be an illogical assumption, since the four entries ab + ab refer to only two questions and thus represent only two agreements. The latter measure implies that

> "coders a and b between them considered that 5 questions were in the Higher Order category. There were thus five possible chances for agreement, though in fact only two agreements were achieved."

This assumption appears logical.

It would have been possible to use a measure more lenient than that employed by White. For example, a negative indication of agreement on a given category would be provided by the formula

$$\frac{2 \text{ (total number of agreements on questions in other categories)}}{\text{total number of questions recorded in other categories}}$$

a measure which, applied to the example above, would give a ratio of 6 to 9. More lenient still would be a rank order correlation, based on the number of questions in a given category recorded for each student by each coder. Having chosen a rigorous measure, the problem was to decide on an acceptable standard of reliability. The work of other researchers gave no consistent guide, partly because reports of reliability trials did not always include the formula used and partly because some researchers based their coding on relatively crude category systems which were likely to/

to be inherently more reliable. The four coders used by Bellack et al (1966), in their content analysis of class lessons, achieved percentage agreements ranging from 84% to 96%; but the coding was based on transcripts and the percentage agreements were achieved after a process of review and arbitration among coders. Claus (1969), in her study on the effect of cueing procedures on the development of higher order questioning skills, claimed rating reliability figures of around 80% overall; but again the work was based on transcripts, and the only distinction involved in coding was between "lower order" (corresponding basically to Bloom's Knowledge category) and "higher order" (representing all the other categories in Bloom's taxonomy). Transcripts were also used by Acheson and Tucker (1971), who reported coder reliability figures ranging from .81 to .89 in an analysis of higher cognitive questioning based on the categories outlined in Minicourse 9. Kelly (1973), who claimed agreement among his three coders "consistently in excess of 90%", also worked from transcripts and separated questions for coding purposes into only two categories, Reasoning and Factual.

By contrast, Berliner (1969), in spite of working from transcripts and distinguishing only two question categories (lower order and higher order), reported that reliability among his three raters was still mediocre after ten hours' training. White (1972) found that agreement among observers using his lesson sampling instrument for question techniques was

- for the main trials 50.7%
- for the supplementary trials 57.6%
- for the field test 57.3%

Bayer (1972), in a multidimensional analysis of primary school lessons, achieved high reliability figures overall among his coders, but low measures of agreement in respect of ratings of cognitive levels of activities (based on the categories of Bloom's taxonomy). The reliability coefficients here were .34, .37, .49, .56, .61.

The/

The most relevant evidence came from the study by Meux and Smith (1964) on logical dimensions of teaching behaviour. Using four coders working in pairs, they determined coefficients of agreement for each of the twenty-five episodes, or verbal transactions, which they had identified. The coefficients were derived from a ratio of actual to possible agreements, and they ranged from .00 to 1.00, with a median of .67. The middle 50 per cent of the coefficients ranged from .62 to .84 with a mean of .68.

The problems in achieving reliability, reported by Meux and Smith, are of interest, in that they anticipated very closely the problems that were to arise for the Jordanhill coders. The main difficulties mentioned were

- the overlapping between the categories
- transactions which did not appear to satisfy the criteria of any one category
- a gradual shading of one category into another ("judges may differ as to what point of the continuum separates one category from the other")
- the high level of inference involved in deciding whether the criteria for certain categories were satisfied

Meux and Smith conclude:

> "At present, we have been unable to develop completely independent categories or suitable criteria by which to eliminate such difficulties, and often it is not clear which need more improvement - the categories or the criteria."

If the reality of a cognitive continuum is accepted, it would follow that the difficulties are built into the nature of any category system; though appropriate choice of categories and precise definitions may certainly reduce the problems.

In the light of the evidence, with particular reference to Meux and Smith, it appeared reasonable for the Jordanhill study to set a target of 70 per cent for inter-coder agreement, with a maximum range of 20 per cent between the highest and lowest levels of agreement for any one category.

After/

After a series of group training sessions, a reliability trial was held in December, 1972. Five-minute extracts from twelve student lessons were coded, with the following results:

TABLE X

Results of December 1972 reliability trial of coding instrument

Question categories	Measures of agreement					
	Coders A - B		Coders A - C		Coders B - C	
Higher Order questions actual agreements possible agreements	5 58	8.62%	47 59	83.05%	5 53	9.43%
Middle Order questions actual agreements possible agreements	82 174	47.13%	93 123	75.61%	92 171	53.80%
Lower Order questions actual agreements possible agreements	84 111	75.68%	94 115	81.74%	93 131	70.99%
Yes/No questions actual agreements possible agreements	23 29	79.31%	23 24	95.83%	24 29	82.76%
One-word answer questions actual agreements possible agreements	52 93	55.91%	78 91	85.71%	55 91	60.44%
No response questions actual agreements possible agreements	27 43	62.79%	32 38	84.21%	27 41	65.85%
No opportunity to respond actual agreements possible agreements	53 74	71.62%	57 66	86.36%	61 73	83.56%
Prompts actual agreements possible agreements	3 22	13.64%	6 22	27.27%	3 6	50.00%
Probes actual agreements possible agreements	2 8	25.00%	7 10	70.00%	2 9	22.22%
Redirections actual agreements possible agreements	nil 2	00.00%	2 2	100.00%	nil 2	00.00%

Total number of questions recorded

	Coder A	Coder B	Coder C
Higher Order	58	5	53
Middle Order	105	151	112
Lower Order	103	118	106
Overall totals	266	274	271

These figures gave clear indications of the problems that needed to be resolved. Coder C was the researcher, coder A was working on a related project with similar ground rules, and coder B. though well versed in research procedures, was not involved in any project concerned with question categories. It appeared that the Ground Rules, while giving an adequate guide to those who shared certain assumptions about the nature and level of questions, were insufficiently explicit to lead an "external" assessor along the same paths of decision-making. In particular, the borderline between Middle Order and Higher Order questions needed sharper definition. The nature of Prompts had obviously caused problems for all three coders and Probing questions had also created difficulties. The occurrence of Redirections was so rare that the figures seemed unlikely to provide a clear indication of reliability.

A revision of the Ground Rules was undertaken, in an attempt to reduce the "grey areas" between Middle Order and Higher Order questions, and between Prompts and Probes. The category of "single-word answer" questions was deleted, since the researcher was unable to arrive at satisfactory definitions of such questions, in terms either of teacher intention or of grammatical construction. The Jordanhill History department had been keen to draw students' attention to the over-use of this type of question, but it was felt that the training given in making optimum use of higher order questions would effectively reduce the number of/

of single-word-answer questions, since most of these came into the lower order category.

The most difficult problem to resolve was the classification of the exploratory variety of questions referred to above; questions which appeared to be characteristic of history teaching and which lay at any point along a continuum extending from simple recall ("What would be the name given to those old merchant ships?") to a higher order prediction involving a fresh act of analysis and synthesis ("What might the pattern of the Second World War have been if the Germans had broken through at El Alamein?"). In between, lay a whole range of questions which appeared to demand the application of general knowledge and common sense, in a process of making "intelligent guesses". The original version of the Ground Rules had included, in the Middle Order category, "questions requiring the pupils to make oredictions or suggestions at a level of common sense and intelligent quesswork, based upon general knowledge of the subject under discussion"; and, in the Higher Order category, "questions requiring the pupils to make inferences or predictions, or develop implications, based upon the material under examination". These two definitions had clearly led the two Stirling coders in different directions, and the revision therefore dropped the phrase "intelligent guesswork" and emphasized the distinction between "applying information, principles or rules already acquired, in order to make suggestions relating to new situations and specific problems" and "framing hypotheses or making predictions, based upon an analysis of given material and involving a logical development of consequences implicit in that material". The danger in sharpening definitions was that one would simply open a gap at a point in the continuum where there had previously been an overlap. Moreover, as the researcher wrote to his two colleagues, "I am still convinced that, if we are to stay close to reality, we must find some distinct category for the type of question, requiring an element of prediction/

prediction, that goes beyond recall of information but does not demand an act of analysis. The difficulty is that the teacher's intention will be governed by his knowledge of or assumptions about the data that the pupils already possess; but the coder has no means of sharing that knowledge, and is accordingly obliged to base his decisions on arbitrary criteria, such as the data provided in the lesson prior to the question or the form of the question itself. It would seem that 'the form of the question, in the absence of evidence pointing in another direction' is the most reliable basis for decision-making and probably no more arbitrary than any other criterion."

After the first reliability trial, the coders reported occasional difficulties in interpreting the audio tape recordings. In particular, it was not always clear whether the teacher was directing follow-up hints (Prompts) or questions (Probes) to the pupil who had been called on to answer the initial question, or to some other pupil; and while it was usually possible to interpret pupil responses, it was sometimes difficult to distinguish one voice from another. These mechanical problems undoubtedly reduced levels of agreement, particularly in relation to Prompts and Probes, but, since there was no practical possibility of using video recording for the main experiment, a switch to video for the reliability trials would have been unjustified.

Training of coders was continued for a further period. Transcripts were made of the lesson excerpts which had provided most disagreement in the December trial, and each problem was identified and discussed. Two subsidiary reliability tests were held, each based on two student lesson excerpts, and these recordings were also transcribed and discussed. The first of these tests, undertaken with the original Ground Rules, revealed persistent difficulties over interpretations of Prompts and Probes, but some improvement in agreeing over Middle Order/Higher Order differentiation. The/

The second test was based on the revision of the Rules, and the results appeared sufficiently satisfactory to warrant a second full-scale reliability trial, using five-minute excerpts from ten student lessons. The results are set out in Table XI, and the mean and range of inter-coder agreements are given in Table XII.

The shift in the mean percentage of overall agreement between December and May was

	1st trial	2nd trial
Coder A - Coder B	42.64%	59.42%
Coder A - Coder C	78.23%	79.04%
Coder B - Coder C	48.73%	69.87%

Clearly, the revision of the Ground Rules and the extra training period had gone some way towards bringing coder B's interpretation of categories nearer to that of coders A and C, but it was disappointing that the target for agreement had not been reached in every instance. The figures which gave particular concern were the comparatively low percentages of agreement (66.15%, 62.73%, 52.65%) for separating questions into the three basic classifications of Higher, Middle and Lower Order. The substantial gains in Higher Order agreements were offset by a sharp drop in Lower Order agreements; indicating that, while the revisions had reduced the grey area between Middle and Higher Order classification, a new grey area between Middle and Lower Order had opened up, though it was difficult to understand why this should have occurred. Agreements over the specialised categories of yes/no, no response and no opportunity to respond questions were reasonably good, but the definitions of Prompting and Redirecting obviously needed further refinement.

Comparisons of percentage agreements for all categories between the first and second trials showed 18 gains and 9 losses. The mean level of agreement between coder A and coder C (who was subsequently to code all the control/

TABLE XI

Results of May 1973 reliability trial of coding instrument

Question categories	Measures of agreement					
	Coders A - B		Coders A - C		Coders B - C	
Higher Order questions						
actual agreements	50	51.04%	69	88.46	56	58.95%
possible agreements	98		78		95	
Middle Order questions						
actual agreements	97	51.03%	123	76.88%	106	60.27%
possible agreements	190		160		176	
Lower Order questions						
actual agreements	34	41.98%	45	61.64%	44	54.32%
possible agreements	81		73		81	
Yes/No questions						
actual agreements	22	81.48%	24	92.31%	23	82.14%
possible agreements	27		26		28	
No response questions						
actual agreements	33	71.74%	37	86.05%	38	82.61%
possible agreements	46		43		46	
No opportunity to respond						
actual agreements	27	65.85%	31	83.78%	31	75.61%
possible agreements	41		37		41	
Prompts						
actual agreements	14	46.67%	20	64.52%	18	72.00%
possible agreements	30		31		25	
Probes						
actual agreements	30	75.00%	32	74.42%	34	82.93%
possible agreements	40		43		41	
Redirections						
actual agreements	5	50.00%	5	83.33%	6	60.00%
possible agreements	10		6		10	

Total number of questions recorded

	Coder A	Coder B	Coder C
Higher Order	72	75	75
Middle Order	144	143	139
Lower Order	54	61	64
Overall Totals	270	279	278

TABLE XII

Mean and range of inter-coder agreements in second reliability trial

Question category	Mean percentage agreement	Range in percentage
Higher Order questions	66.15	37.42
Middle Order questions	62.73	25.85
Lower Order questions	52.65	19.66
Yes/No questions	85.31	10.83
No response questions	80.13	14.31
No opportunity to respond	75.08	17.93
Prompts	61.06	25.33
Probes	77.45	7.51
Redirections	64.44	33.33

control and experimental class lessons) was satisfactory and consistent in both trials. Nevertheless, more work needed to be done before the coding instrument could be thought reliable in the hands of a researcher working on his own. Unfortunately, it was not possible to hold a third reliability trial, because coder A left the country immediately after the second trial. It was therefore decided to carry out an analysis of disagreements, concentrating on coders A and B, since the level of A - B agreement in the May trial was lower, in every instance except one, than either the A - C or the B - C level. An analysis was also made of coder C's disagreements in the Lower Order question category, where his level of agreement with both A and B had fallen below 70%.

The breakdown of coder A - B disagreements was

Higher/

Higher Order/Middle Order

Total number of disagreements	32
A scored Higher Order	11
B scored Higher Order	21

Lower Order/Middle Order

Total number of disagreements	30
A scored Middle Order	15
B scored Middle Order	15

Higher Order/Lower Order

Total number of disagreements	2
A scored Higher Order in both cases	
Total	64

These totals excluded disagreements arising from failure by one or other coder (noted on the coding sheet) to understand what was said. They also excluded those instances where one coder failed to note a question recorded by the other coder.

In addition, there were 16 instances of disagreement over Prompts. These were tabled separately, since not all Prompts were in question form.

The analysis brought to light the main reasons for disagreement over the three basic question categories. For example, eleven disagreements arose from indecision over an appropriate category for questions beginning "Do you think ...?" or "Would you agree ...?". Such questions often appeared to demand no more than a yes/no response, and coders were uncertain whether an effort of analysis or simply an unsupported opinion was being called for. Indecision over the appropriate category for questions with "What would...?" stems still accounted for sixteen disagreements. It was difficult/

difficult to be sure whether disagreements over Prompts arose from deliberate decisions or from oversights, but there were indications that in some instances coders were failing to note that the Prompt had to support the initial question by providing some kind of a clue or by rephrasing the question in a simpler form.

A further set of revisions and additions was written into the Ground Rules, covering 62 of the 64 identified disagreements. Thereafter, it was considered reasonable to assume reliability for the coding system. The considerable gain in the A - B and B - C mean levels of agreement between December and May suggested that, given further training in what was undoubtedly a highly complex instrument, all three coders would have gone on to achieve the 70 per cent target of agreement. Meantime, the high levels of agreement between coders A and C indicated that the instrument was already reliable in the hands of researchers working closely together in a similar field. The long period taken to achieve a reasonable measure of agreement perhaps argues that a cruder, more arbitrary system of categories would have been more practicable; but it must be remembered that the coders were all engaged on other tasks and that the training programme was therefore necessarily spasmodic. Moreover, if the coding was to be valid as well as reliable, it seemed necessary to develop an instrument which would sensitively reflect "the extraordinarily complex interactions contained in the classroom discourse" (Bellack, 1966). The problems encountered in formulating the Jordanhill Ground Rules derived not only from this inherent complexity but also from the need to match the definitions of skills to the objectives of the History department. There were difficulties of agreement which could have been eased by a change of definition, but the result would have been a coding of behaviours which diverged from training procedures. It was important not to purchase reliability at the expense of relevance.

6. Results of the main experiment

The data from the recorded classroom lessons in terms 1 and 3 of the main experiment, on which all the quantitative measures were based, are set out in Tables XIIIa - XIIIh. In each half term, students 1 to 9 are numbered in the order of their ranking by tutors for their term 1 performance on the selected questioning skills. Only nine students are listed in each group, since, in each half of term 2, illness forced one of the experimental group to drop out of the microteaching sessions. Since the comparisons were based on matched pairs, the two control group partners had also to be omitted from the analysis.

The symbols used throughout the Tables in this chapter are defined as follows:

HOQ	-	Higher Order question
MOQ	-	Middle Order question
LOQ	-	Lower Order question
P	-	Prompt
XP	-	Prompt following a question which obtained no response
Pr	-	Probe (of initial response)
R	-	Redirection (of initial response to other pupils, for comment and discussion)
y	-	Question in a form which appeared to demand no more than a 'yes' or 'no' response
x	-	Question which attracted no response
n	-	Question where no opportunity was given to respond

Ideally, the classroom lessons should have been coded by someone who had no knowledge of which students were in the experimental and control groups, but in fact all the coding had to be undertaken by the researcher. Fortunately, control and experimental group student lessons were randomly mingled on each tape, and the tapes were selected for coding in random order, to avoid any term 1-term 3 bias. The coding was carried out several months after the microteaching sessions, and the researcher deliberately refrained from checking the names of the students in the experiment and control lists until all the analysis was complete.

Table XIIIa

Experimental group data: first half term 1

STUDENTS	1	2	3	4	5	6	7	8	9	TOTALS
Total Q excl. n	43	34	17	52	57	32	34	50	30	349
Total HOQ	-	6	4	16	5	2	2	15	-	50
% HOQ / Q	0.0	17.6	23.5	30.8	8.8	6.3	5.9	30.0	0.0	M=13.66
Total MOQ	16	27	9	22	31	19	16	12	15	167
% MOQ / Q	37.2	79.4	52.9	42.3	54.4	59.4	47.1	24.0	50.0	M=49.63
Total LOQ	27	1	4	14	21	11	16	23	15	132
% LOQ / Q	62.8	2.9	23.5	26.9	36.8	34.4	47.1	46.0	50.0	M=36.71
Total PROMPTS	1	2	1	9	2	3	4	1	7	30
Total XP	1	-	1	6	2	3	3	1	7	24
% XP / X	50.0	0.0	33.3	50.0	22.2	75.0	60.0	12.5	77.8	M=42.31
Total PROBES	1	1	4	3	1	1	1	5	-	17
% Pr / Q	2.3	2.9	23.5	5.8	1.8	3.1	2.9	10.0	0.0	M=5.81
Total REDIRECTIONS	2	-	-	-	-	-	-	-	-	2
% R / Q	4.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	M=0.52
Total y	3	-	2	1	2	2	1	2	1	14
% y / Q	7.0	0.0	11.8	1.9	3.5	6.3	2.9	4.0	3.3	M=4.52
Total X	2	1	3	12	9	4	5	8	9	53
% X / Q	4.7	2.9	17.6	23.1	15.8	12.5	14.7	16.0	30.0	M=15.26
Total HO X	-	-	-	3	4	-	-	1	-	8
% HO X / HOQ	0.0	0.0	0.0	18.8	80.0	0.0	0.0	6.7	0.0	M=11.72
Total n	1	3	8	10	14	7	2	5	-	50
Total Q incl. n	44	37	25	62	71	39	36	55	30	399
% n / Q incl. n	2.3	8.1	32.0	16.1	19.7	17.9	5.6	9.1	0.0	M=12.31

Table XIIIb

Control group data: first half term 1

STUDENTS	1	2	3	4	5	6	7	8	9	TOTALS
Total Q excl. n	37	42	30	64	47	46	23	32	42	363
Total HOQ	6	–	1	8	5	–	–	6	1	27
% HOQ / Q	16.2	0.0	3.3	12.5	10.6	0.0	0.0	18.8	2.4	M=7.09
Total MOQ	23	25	14	31	15	9	15	24	28	184
% MOQ / Q	62.2	59.5	46.7	48.4	31.9	19.6	65.2	75.0	66.7	M=52.8
Total LOQ	8	17	15	25	27	37	8	2	13	152
% LOQ / Q	21.6	40.5	50.0	39.1	57.4	80.4	34.8	6.3	31.0	M=40.12
Total PROMPTS	2	6	1	9	1	1	–	1	3	24
Total XP	2	6	1	8	1	–	–	1	3	22
% XP / X	50.0	66.7	33.3	88.9	16.7	0.0	0.0	100.0	33.3	M=43.21
Total PROBES	1	3	1	3	1	–	–	1	2	12
% Pr / Q	2.7	7.1	3.3	4.7	2.1	0.0	0.0	3.1	4.8	M=3.09
Total REDIRECTIONS	–	–	–	–	–	–	–	–	–	0
% R / Q	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	M=0.0
Total y	13	–	1	1	7	3	–	1	5	31
% y / Q	35.1	0.0	3.3	1.6	14.9	6.5	0.0	3.1	11.9	M=8.49
Total X	4	9	3	9	6	1	–	1	9	42
% X / Q	10.8	21.4	10.0	14.1	12.8	2.2	0.0	3.1	21.4	M=10.64
Total HO X	1	–	–	–	–	–	–	–	–	1
% HO X / HOQ	16.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	M=1.86
Total n	4	10	3	6	10	15	1	2	10	61
Total Q incl. n	41	52	33	70	57	61	24	34	52	424
% n / Q incl. n	9.8	19.2	9.1	8.6	17.5	24.6	4.2	5.9	19.2	M=13.12

Table XIIIc

Experimental group data: second half term 1

STUDENTS	1	2	3	4	5	6	7	8	9	TOTALS
Total Q excl. n	50	32	29	44	49	29	73	43	22	371
Total HOQ	1	3	3	4	4	1	5	9	12	42
% HOQ / Q	2.0	9.4	10.3	9.1	8.2	3.4	6.8	20.9	54.5	M=13.84
Total MOQ	27	11	17	16	18	22	22	19	6	158
% MOQ / Q	54.0	34.4	58.6	36.4	36.7	75.9	30.1	44.2	27.3	M=44.18
Total LOQ	22	18	9	24	27	6	46	15	4	171
% LOQ / Q	44.0	56.3	31.0	54.5	55.1	20.7	63.0	34.9	18.2	M=41.97
Total PROMPTS	2	4	4	4	5	1	6	2	-	28
Total XP	2	4	4	3	5	1	5	2	-	26
% XP / X	18.2	44.4	80.0	75.0	41.7	50.0	50.0	28.6	0.0	M=43.1
Total PROBES	1	-	5	7	3	4	7	4	3	34
% Pr / Q	2.0	0.0	17.2	15.9	6.1	13.8	9.6	9.3	13.6	M=9.72
Total REDIRECTIONS	-	-	-	1	-	-	-	-	-	1
% R / Q	0.0	0.0	0.0	2.3	0.0	0.0	0.0	0.0	0.0	M=0.26
Total y	8	2	5	1	6	3	2	3	-	30
% y / Q	16.0	6.3	17.2	2.3	12.2	10.3	2.7	7.0	0.0	M=8.22
Total X	11	9	5	4	12	2	10	7	-	60
% X / Q	22.0	28.1	17.2	9.1	24.5	6.9	13.7	16.3	0.0	M=15.31
Total HO X	-	2	2	1	1	-	1	1	-	8
% HO X / HOQ	0.0	66.7	66.7	25.0	25.0	0.0	20.0	11.1	0.0	M=23.83
Total n	15	2	6	2	14	11	4	1	1	56
Total Q incl. n	65	34	35	46	63	40	77	44	23	427
% n / Q incl. n	23.1	5.9	17.1	4.3	22.2	27.5	5.2	2.3	4.3	M=12.43

Table XIIId

Control group data: second half term 1

STUDENTS	1	2	3	4	5	6	7	8	9	TOTALS
Total Q excl. n	59	39	32	36	33	24	34	35	23	315
Total HOQ	8	2	4	4	7	8	2	3	7	45
% HOQ/Q	13.6	5.1	12.5	11.1	21.2	33.3	5.9	8.6	30.4	M=15.74
Total MOQ	20	16	20	10	15	13	10	10	13	127
% MOQ/Q	33.9	41.0	62.5	27.8	45.5	54.2	29.4	28.6	56.5	M=42.16
Total LOQ	31	21	8	22	11	3	22	22	3	143
% LOQ/Q	52.5	53.8	25.0	61.1	33.3	12.5	64.7	62.9	13.0	M=42.09
Total PROMPTS	5	2	3	-	1	1	1	1	2	16
Total XP	3	2	3	-	-	1	1	1	1	12
% XP/X	50.0	20.0	60.0	0.0	0.0	50.0	50.0	50.0	50.0	M=36.67
Total PROBES	5	-	-	-	2	-	1	3	2	13
% Pr/Q	8.5	0.0	0.0	0.0	6.1	0.0	2.9	8.6	8.7	M=3.87
Total REDIRECTIONS	-	-	2	-	-	-	-	-	-	2
% R/Q	0.0	0.0	6.3	0.0	0.0	0.0	0.0	0.0	0.0	M=0.7
Total y	5	2	5	-	8	5	3	2	1	31
% y/Q	8.5	5.1	15.6	0.0	24.2	20.8	8.8	5.7	4.3	M=10.33
Total X	6	10	5	1	-	2	2	2	2	30
% X/Q	10.2	25.6	15.6	2.8	0.0	8.3	5.9	5.7	8.7	M=9.2
Total HO X	2	1	-	-	-	1	-	-	1	5
% HO X/HOQ	25.0	50.0	0.0	0.0	0.0	12.5	0.0	0.0	14.3	M=11.31
Total n	5	2	4	-	6	1	9	4	2	33
Total Q incl. n	64	41	36	36	39	25	43	39	25	348
% n/Q incl. n	7.8	4.9	11.1	0.0	15.4	4.0	20.9	10.3	8.0	M=9.16

Table XIIIe

Experimental group data: first half term 3

STUDENTS	1	2	3	4	5	6	7	8	9	TOTALS
Total Q excl. n	30	49	42	52	35	38	39	52	46	383
Total HOQ	1	10	19	19	15	17	18	22	12	133
% HOQ / Q	3.3	20.4	45.2	36.5	42.9	44.7	46.2	42.3	26.1	M=34.18
Total MOQ	14	35	22	24	15	19	11	21	34	195
% MOQ / Q	46.7	71.4	52.4	46.2	42.9	50.0	28.2	40.4	73.9	M=50.23
Total LOQ	15	4	1	9	5	2	10	9	-	55
% LOQ / Q	50.0	8.2	2.4	17.3	14.3	5.3	25.6	17.3	0.0	M=15.6
Total PROMPTS	1	1	1	7	2	2	3	3	2	22
Total XP	1	1	-	7	2	-	2	3	2	18
% XP / X	50.0	100.0	0.0	70.0	66.7	0.0	66.7	42.9	100.0	M=55.14
Total PROBES	4	7	7	7	2	4	3	9	6	49
% Pr / Q	13.3	14.3	16.7	13.5	5.7	10.5	7.7	17.3	13.0	M=12.44
Total REDIRECTIONS	-	2	4	-	-	1	4	3	-	14
% R / Q	0.0	4.1	9.5	0.0	0.0	2.6	10.3	5.8	0.0	M=3.59
Total y	2	1	5	3	-	5	3	4	1	24
% y / Q	6.7	2.0	11.9	5.8	0.0	13.2	7.7	7.7	2.2	M=6.36
Total X	2	1	-	10	3	5	3	7	2	33
% X / Q	6.7	2.0	0.0	19.2	8.6	13.2	7.7	13.5	4.3	M=8.36
Total HO X	-	1	-	4	-	3	2	5	1	16
% HO X / HOQ	0.0	100.0	0.0	21.1	0.0	17.6	11.1	22.7	8.3	M=20.09
Total n	3	1	7	6	10	4	8	-	3	42
Total Q incl. n	33	50	49	58	45	42	47	52	49	425
% n / Q incl. n	9.1	2.0	14.3	10.3	22.2	9.5	17.0	0.0	6.1	M=10.06

Table XIIIf

Control group data: first half term 3

STUDENTS	1	2	3	4	5	6	7	8	9	TOTALS
Total Q excl. n	30	68	26	58	46	44	20	33	40	365
Total HOQ	-	5	2	3	3	3	4	5	2	27
% HOQ / Q	0.0	7.4	7.7	5.2	6.5	6.8	20.0	15.2	5.0	M=8.2
Total MOQ	5	17	12	17	30	33	10	23	17	164
% MOQ / Q	16.7	25.0	46.2	29.3	65.2	75.0	50.0	69.7	42.5	M=46.62
Total LOQ	25	46	12	38	13	8	6	5	21	174
% LOQ / Q	83.3	67.6	46.2	65.5	28.3	18.2	30.0	15.2	52.5	M=45.2
Total PROMPTS	4	12	3	7	2	2	-	3	7	40
Total XP	3	10	2	6	2	2	-	3	6	34
% XP / X	60.0	35.7	50.0	75.0	33.3	66.7	0.0	60.0	100.0	M=53.41
Total PROBES	1	5	1	2	3	8	2	-	-	22
% Pr / Q	3.3	7.4	3.8	3.4	6.5	18.2	10.0	0.0	0.0	M=5.84
Total REDIRECTIONS	-	-	-	-	-	-	-	1	-	1
% R / Q	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.0	0.0	M=0.33
Total y	4	-	1	3	2	9	1	4	-	24
% y / Q	13.3	0.0	3.8	5.2	4.3	20.5	5.0	12.1	0.0	M=7.13
Total X	5	28	4	8	6	3	-	5	6	65
% X / Q	16.7	41.2	15.4	13.8	13.0	6.8	0.0	15.2	15.0	M=15.23
Total HO X	-	2	-	-	1	-	-	2	-	5
% HO X / HOQ	0.0	40.0	0.0	0.0	33.3	0.0	0.0	40.0	0.0	M=12.59
Total n	3	14	2	6	16	19	3	10	8	81
Total Q incl. n	33	82	28	64	62	63	23	43	48	446
% n / Q incl. n	9.1	17.1	7.1	9.4	25.8	30.2	13.0	23.3	16.7	M=16.86

Table XIIIg

Experimental group data: second half term 3

STUDENTS	1	2	3	4	5	6	7	8	9	TOTALS
Total Q excl. n	54	38	32	21	41	45	34	45	53	363
Total HOQ	4	5	17	8	18	6	13	7	11	89
% HOQ/Q	7.4	13.2	53.1	38.1	43.9	13.3	38.2	15.6	20.8	M=27.07
Total MOQ	34	22	15	9	19	26	15	31	36	207
% MOQ/Q	63.0	57.9	46.9	42.9	46.3	57.8	44.1	68.9	67.9	M=55.08
Total LOQ	16	11	-	4	4	13	6	7	6	67
% LOQ/Q	29.6	28.9	0.0	19.0	9.8	28.9	17.6	15.6	11.3	M=17.86
Total PROMPTS	3	4	-	1	2	6	1	8	8	33
Total XP	2	3	-	1	2	5	1	8	8	30
% XP/X	25.0	75.0	0.0	50.0	33.3	50.0	33.3	88.9	80.0	M=48.39
Total PROBES	2	3	9	6	2	5	3	3	4	37
% Pr/Q	3.7	7.9	28.1	28.6	4.9	11.1	8.8	6.7	7.5	M=11.92
Total REDIRECTIONS	-	-	1	-	-	3	-	-	-	4
% R/Q	0.0	0.0	3.1	0.0	0.0	6.7	0.0	0.0	0.0	M=1.09
Total y	9	1	9	1	1	3	-	3	4	31
% y/Q	16.7	2.6	28.1	4.8	2.4	6.7	0.0	6.7	7.5	M=8.39
Total X	8	4	3	2	6	10	3	9	10	55
% X/Q	14.8	10.5	9.4	9.5	14.6	22.2	8.8	20.0	18.9	M=14.3
Total HO X	1	-	-	2	5	1	-	2	2	13
% HO X/HOQ	25.0	0.0	0.0	25.0	27.8	16.7	0.0	28.6	18.2	M=15.7
Total n	10	1	2	1	13	18	3	3	5	56
Total Q incl. n	64	39	34	22	54	63	37	48	58	419
% n/Q incl. n	15.6	2.6	5.9	4.5	24.1	28.6	8.1	6.3	8.6	M=11.59

Table XIIIh

Control group data: second half term 3

STUDENTS	1	2	3	4	5	6	7	8	9	TOTALS
Total Q excl. n	47	45	40	39	30	39	34	38	30	342
Total HOQ	3	17	3	3	10	7	3	4	3	53
% HOQ / Q	6.4	37.8	7.5	7.7	33.3	17.9	8.8	10.5	10.0	M=15.54
Total MOQ	21	16	23	34	6	25	11	8	10	154
% MOQ / Q	44.7	35.6	57.5	87.2	20.0	64.1	32.4	21.1	33.3	M=43.99
Total LOQ	23	12	14	2	14	7	20	26	17	135
% LOQ / Q	48.9	26.7	35.0	5.1	46.7	17.9	58.8	68.4	56.7	M=40.47
Total PROMPTS	2	5	1	1	1	-	1	-	5	16
Total XP	2	5	1	-	1	-	-	-	4	13
% XP / X	25.0	41.7	100.0	0.0	14.3	0.0	0.0	0.0	50.0	M=25.67
Total PROBES	5	-	3	1	5	3	3	6	2	28
% Pr / Q	10.6	0.0	7.5	2.6	16.7	7.7	8.8	15.8	6.7	M=8.49
Total REDIRECTIONS	-	-	-	-	4	-	-	-	-	4
% R / Q	0.0	0.0	0.0	0.0	13.3	0.0	0.0	0.0	0.0	M=1.48
Total y	4	8	2	2	21	5	6	7	-	55
% y / Q	8.5	17.8	5.0	5.1	70.0	12.8	17.6	18.4	0.0	M=17.24
Total X	8	12	1	2	7	1	2	1	8	42
% X / Q	17.0	26.7	2.5	5.1	23.3	2.6	5.9	2.6	26.7	M=12.49
Total HO X	-	5	-	-	2	-	-	-	3	10
% HO X / HOQ	0.0	29.4	0.0	0.0	20.0	0.0	0.0	0.0	100.0	M=16.6
Total n	4	5	2	1	8	9	18	5	13	65
Total Q incl. n	51	50	42	40	38	48	52	43	43	407
% n / Q incl. n	7.8	10.0	4.8	2.5	21.1	18.8	34.6	11.6	30.2	M=15.71

A. Measurements based on coding of recorded lessons.

(i) t-tests to establish significance of differences between term 1 and term 3 performances in selected skills.

To determine the values of t, the formula for matched subject designs was used:

$$t = \bar{d} \div \sqrt{\frac{\sum d^2 - \frac{(\sum d)^2}{n}}{n\,(n-1)}}$$

where $\bar{d}$ = the mean of the differences (d) between term 1 and term 3 performance

n = number of students in each half-term sample

The degrees of freedom were (n-1). t values were completed as follows:

Table XIV

Term 3 - term 1 t values for performance of experimental and control groups in selected skills

SKILLS	EXPERIMENTAL				CONTROL			
	1st half term	Significance	2nd half term	Significance	1st half term	Significance	2nd half term	Significance
		P<		P<		P<		P<
$\frac{HOQ}{Q}$	4.076	.005	1.643		0.321		-0.032	
$\frac{LOQ}{Q}$	-4.157	.005	-4.053	.005	0.425		-0.175	
$\frac{XP}{X}$	1.077		-0.278		0.809		-1.262	
$\frac{Pr}{Q}$	3.389	.005	0.991		1.145		3.346	.02
$\frac{y}{Q}$	1.669		0.080		-0.359		1.226	
$\frac{x}{Q}$	-2.066	.05	-0.254		1.786		1.146	
$\frac{HOX}{HOQ}$	-0.249		-0.682		-		-	
$\frac{n}{Q \text{ incl. } n}$	-0.718		-0.468		1.645		2.235	

In connection with Table XIV, it should be noted that

- tests of significance are 1-tailed for the experimental group and 2-tailed for the control group

- t-values for $\frac{MOQ}{Q}$ were not computed. Since the aim of the microteaching/

microteaching programme was to encourage students to raise the cognitive level of their questioning without necessarily increasing the total number of their questions, it was recognized that changes in the proportion of Higher, Middle and Lower Order questions were likely to exhibit a degree of interdependence. It was therefore not expected that the experimental groups would show any significant change in the number of their Middle Order questions. The effectiveness of the training would be revealed by a proportionate reduction of Lower Order questions and a proportionate increase in Higher Order questions.

- t-values for $\frac{R}{Q}$ were not computed, since the majority of the students asked no questions in this category. Table XIII shows, however, that, between term 1 and term 3, the experimental group increased the total of Redirections by a factor of 6 (3R in term - : 18R in term 3), while the control group Redirections increased by a factor of 2.5 (2R in term 1 : 5R in term 3).

- t-values for control group term 3 - term 1 differences in $\frac{HOX}{HOQ}$ were not computed, because the high proportion of nil-nil comparisons would have rendered any statistical procedure meaningless. However, the sums of such percentage differences as did occur were positive for the control group in each half term. (The corresponding figures for the experimental group were negative.)

- to favour term 3 in relation to term 1 performances, the t-values should be

positive/

positive for $\frac{HOQ}{Q}$, $\frac{XP}{x}$, $\frac{Pr}{Q}$

negative for $\frac{LOQ}{Q}$, $\frac{y}{Q}$, $\frac{x}{Q}$, $\frac{HOX}{HOQ}$, $\frac{n}{Q \text{ incl. } n}$

(ii) t-tests to establish significance of differences between experimental and control groups in term 1 and term 3.

The same formula was used to determine the values of t, with d standing for the difference between experimental and control group performance.

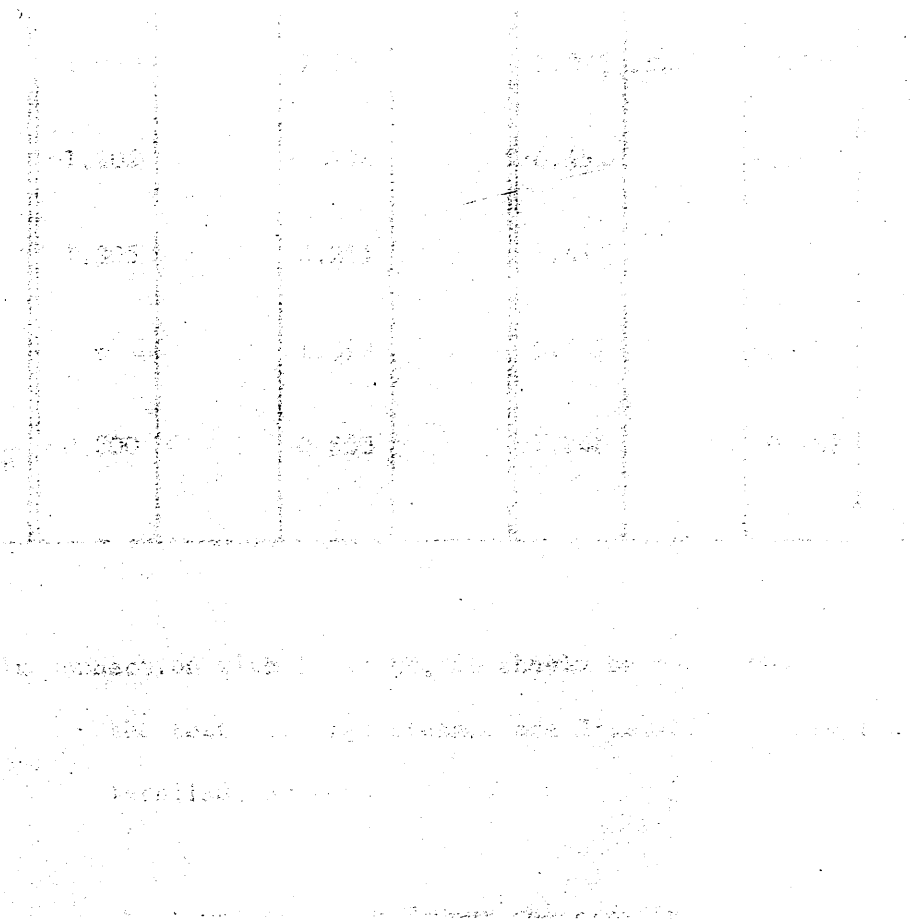


Table XV

Experimental-control t values for performance in selected skills, term 1 and term 3

SKILLS	EXPERIMENTAL-CONTROL				EXPERIMENTAL-CONTROL			
	1st half Term 1	Significance	2nd half Term 1	Significance	1st half Term 3	Significance	2nd half Term 3	Significance
		P<		P<		P<		P<
HOQ / Q	1.653		-0.367		6.588	.0005	1.640	
LOQ / Q	-0.314		-0.027		-3.905	.005	-2.655	.02
XP / X	-0.443		0.444		0.113		1.308	
Pr / Q	1.090		2.134		2.372	.02	0.796	
y / Q	-1.102		-0.976		-0.454		-1.038	
x / Q	1.305		2.211		-1.410		0.456	
HOX / HOQ	-		1.363		-0.369		-0.075	
n / Q incl. n	-0.200		0.833		-1.846		-0.982	

In connection with Table XV, it should be noted that

- the tests of significance are 2-tailed for term 1 and 1-tailed for term 3.

- the t values which favour the experimental group in term 3 are

positive for $\frac{HOQ}{Q}$, $\frac{XP}{X}$, $\frac{Pr}{Q}$

negative/

negative for $\frac{LOQ}{Q}$, $\frac{y}{Q}$, $\frac{x}{Q}$, $\frac{HOX}{HOQ}$, $\frac{n}{Q \text{ incl. } n}$

(iii) t-tests to evaluate significance of differences between gain scores, term 1 to term 3, by experimental and control group.

Again, the same formula was used in determining the value of t, with d standing for the difference between gains.

Table XVI

Experimental-control t values for differences in gain scores, term 1 to term 3

First half term (1-tailed tests)

SKILLS	Difference favouring			
	EXPERIMENTAL		CONTROL	
	t-value	Sig.	t-value	Sig.
		P<		P<
$\frac{HOQ}{Q}$	4.829	.005		
$\frac{LOQ}{Q}$	-2.293	.05		
$\frac{XP}{x}$	0.291			
$\frac{Pr}{Q}$	1.164			
$\frac{y}{Q}$			1.086	
$\frac{x}{Q}$	-4.350	.005		
$\frac{HOX}{HOQ}$	_(exptl superior)			
$\frac{n}{Q \text{ incl. } n}$	-1.567			

Table XVI (Contd.)

Second half term (1-tailed tests)

SKILLS	Difference favouring			
	EXPERIMENTAL		CONTROL	
	t-value	Sig.	t-value	Sig.
		P		P
HOQ / Q	1.619			
LOQ / Q	-2.494	.02		
XP / x	0.522			
Pr / Q			-0.894	
y / Q	-0.919			
x / Q	-0.961			
HOX / HOQ	-0.967			
n / Q incl. n	-3.568	.005		

(iv) Graphical comparison of group performance on each of the selected skills, term 1 to term 3.

For each skill category there were 72 scores (36 for term 1 and 36 for term 3). Means and standard deviations were computed for each group of 72 scores, and each score was expressed in terms of its/

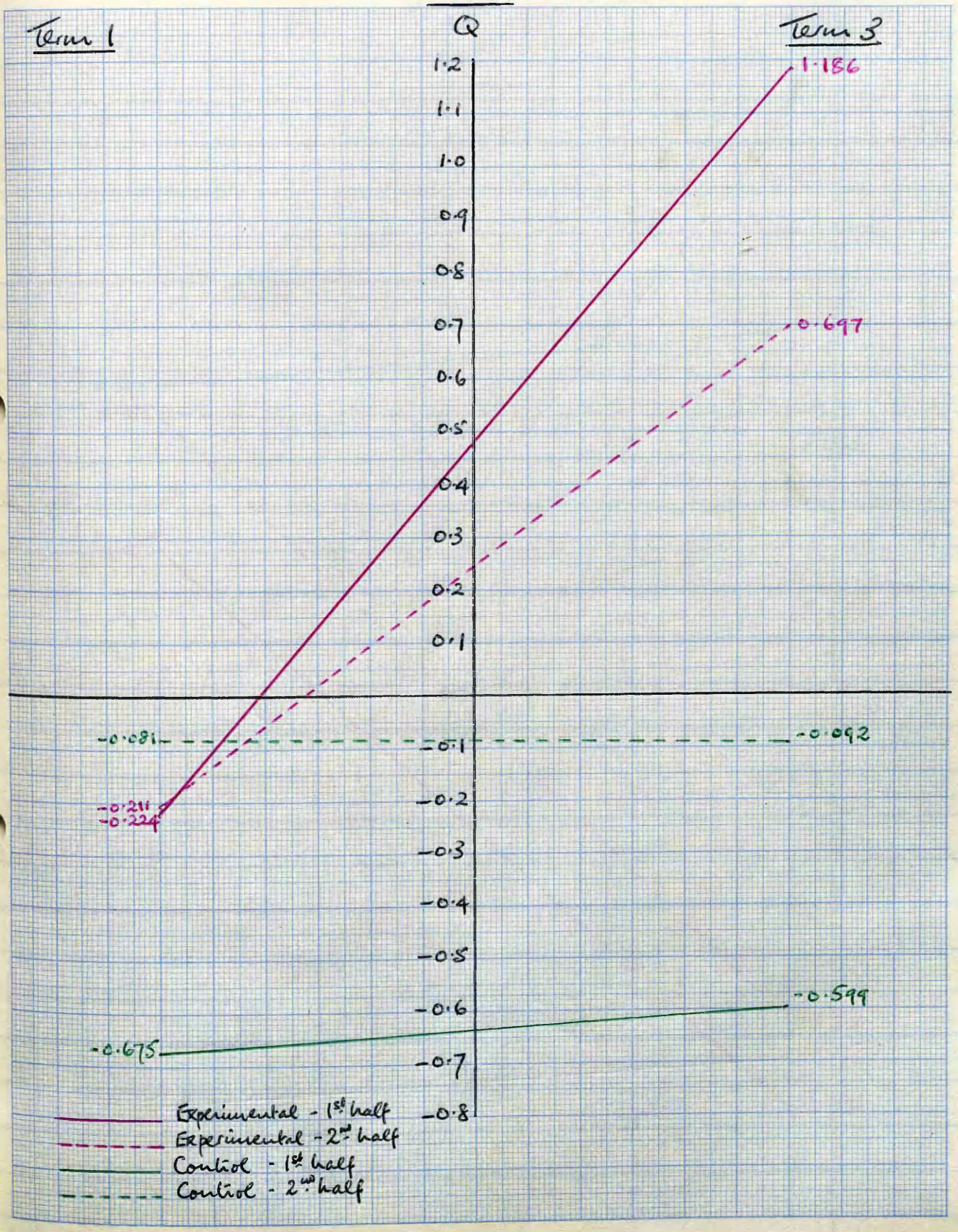
its deviation from the mean. It was then possible to derive standardized scores from these raw scores, using the formula

$$z = \frac{x}{\sigma}$$

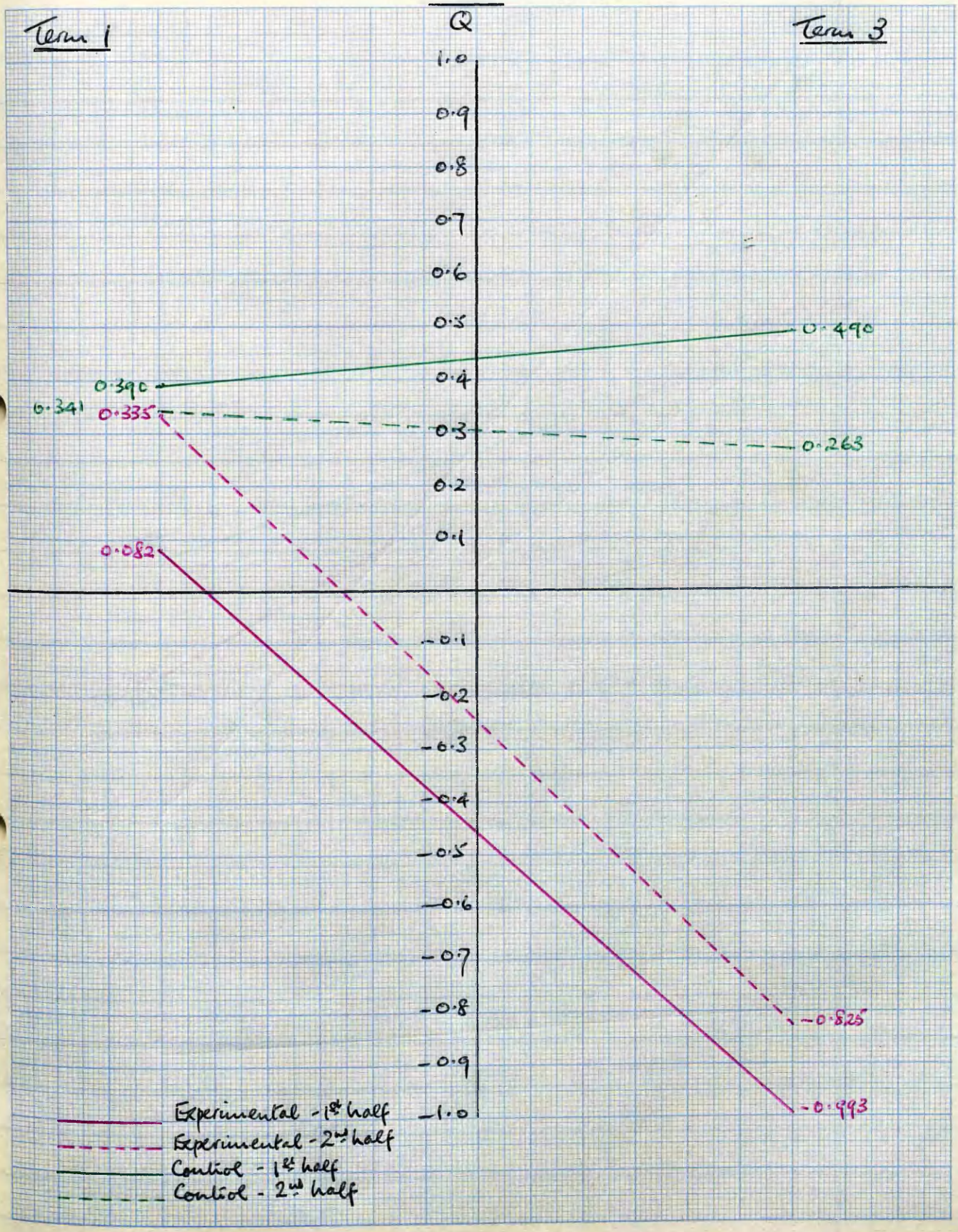
where z = standardized score; x = deviation of raw score from its mean; σ = standard deviation of each group of raw scores.

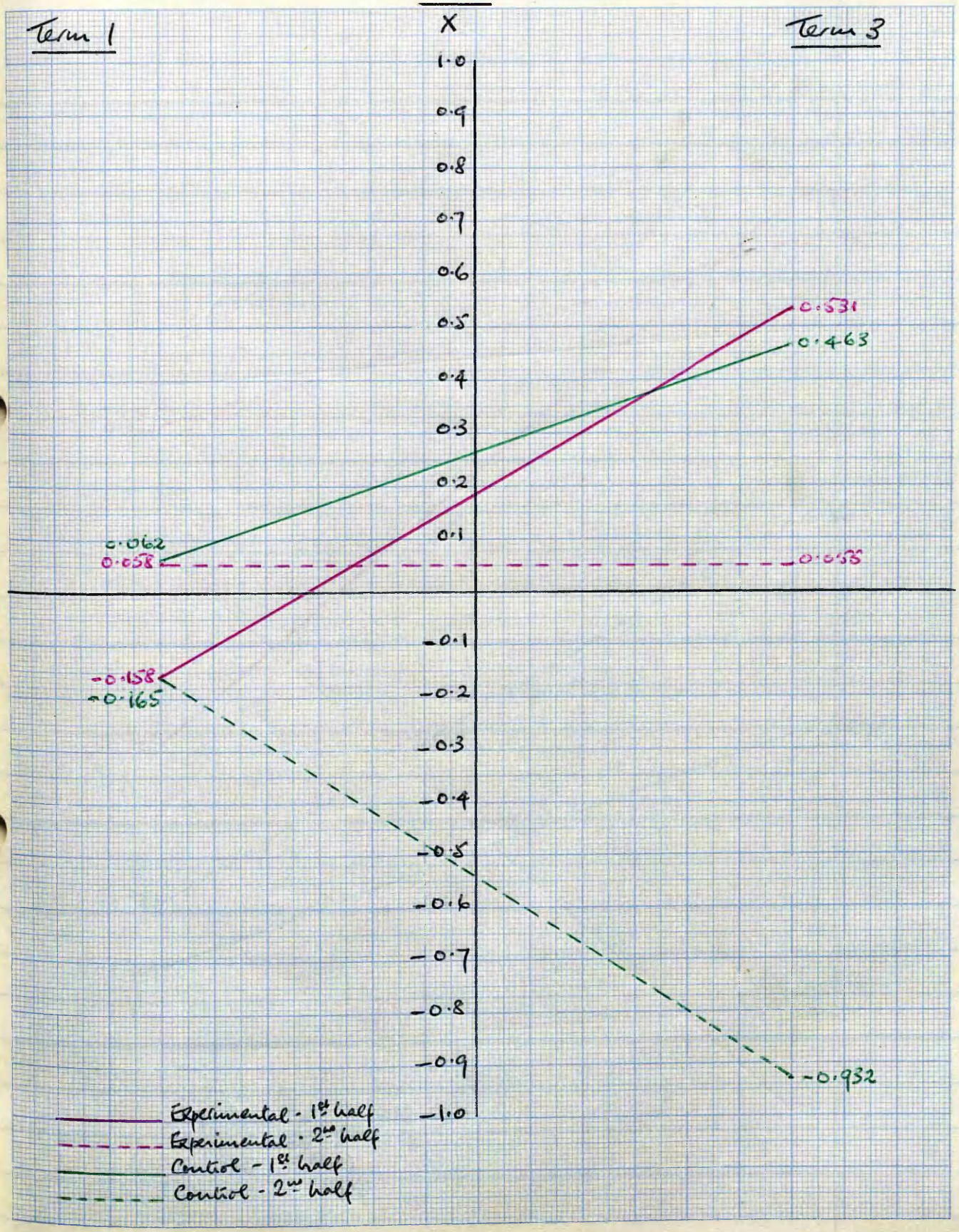
Means of the standardized scores were then computed for each group of students (experimental and control, first and second half terms). This procedure enabled graphical comparisons to be made, about a zero mean, of group performances on each of the selected skills in term 1 and term 3.

HO Q


Term 1
Q
Term 3
1·2
1·1
1·0
0·9
0·8
0·7
0·6
0·5
0·4
0·3
0·2
0·1
−0·1
−0·2
−0·3
−0·4
−0·5
−0·6
−0·7
−0·8
1·186
0·697
−0·081
−0·092
−0·211
−0·224
−0·599
−0·675
Experimental - 1st half
Experimental - 2nd half
Control - 1st half
Control - 2nd half

LOQ


Term 1
Q
Term 3
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
-0.1
-0.2
-0.3
-0.4
-0.5
-0.6
-0.7
-0.8
-0.9
-1.0
0.490
0.390
0.341
0.335
0.263
0.082
-0.825
-0.993
Experimental - 1st half
Experimental - 2nd half
Control - 1st half
Control - 2nd half

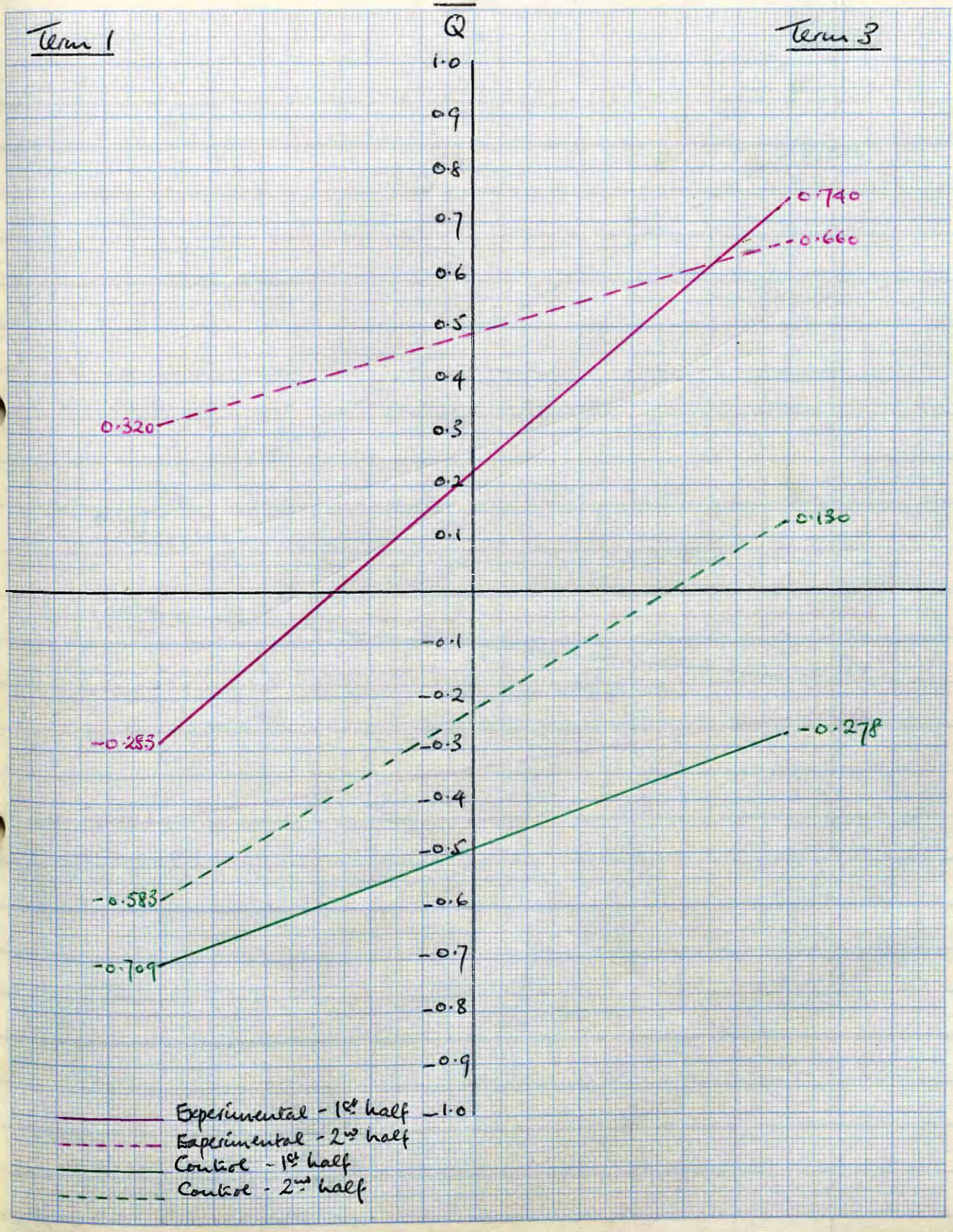

xP
X
Term 1
Term 3
1·0
0·9
0·8
0·7
0·6
0·5
0·4
0·3
0·2
0·1
-0·1
-0·2
-0·3
-0·4
-0·5
-0·6
-0·7
-0·8
-0·9
-1·0
0·531
0·463
0·062
0·058
0·055
-0·158
-0·165
-0·932
Experimental - 1st half
Experimental - 2nd half
Control - 1st half
Control - 2nd half

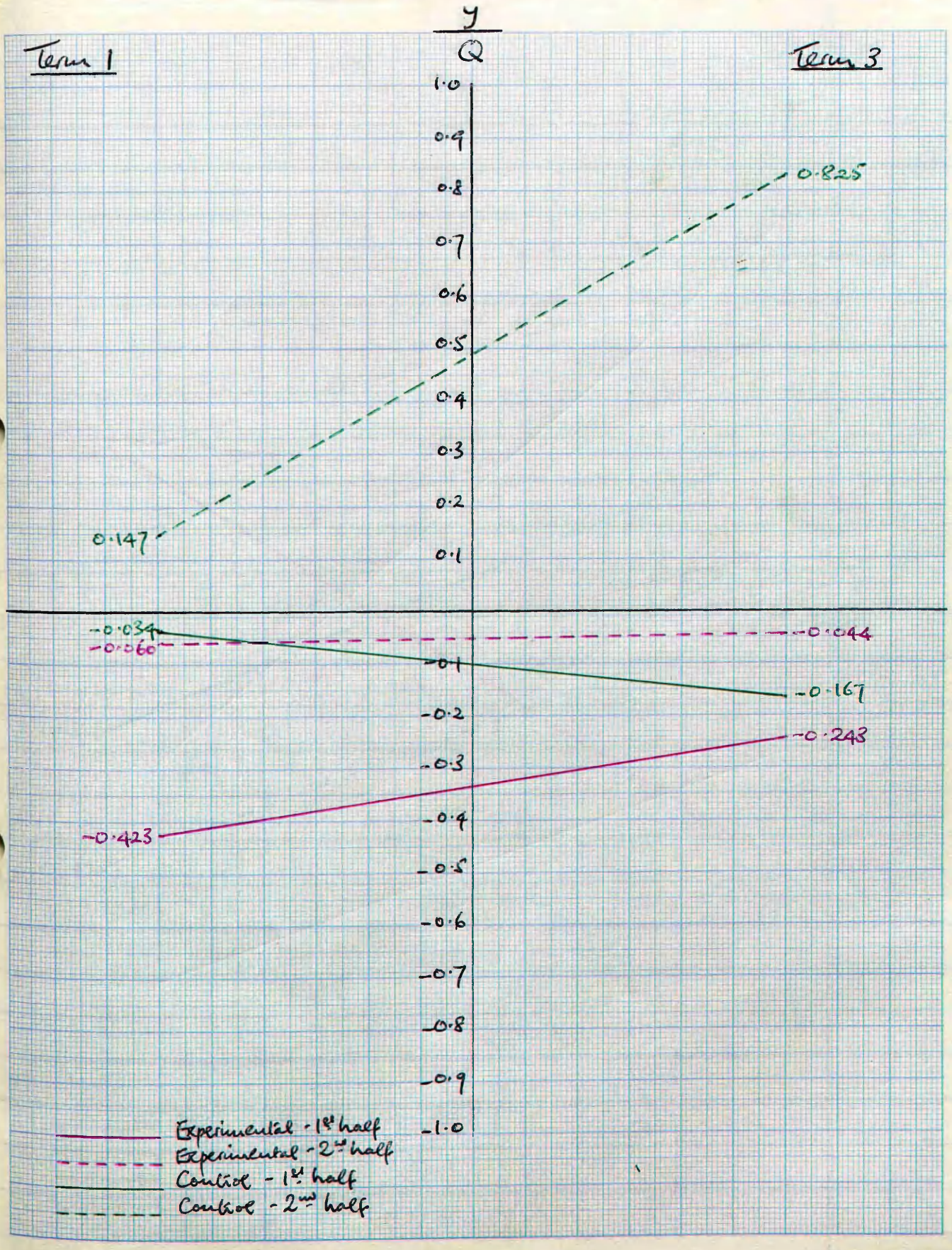
Pr

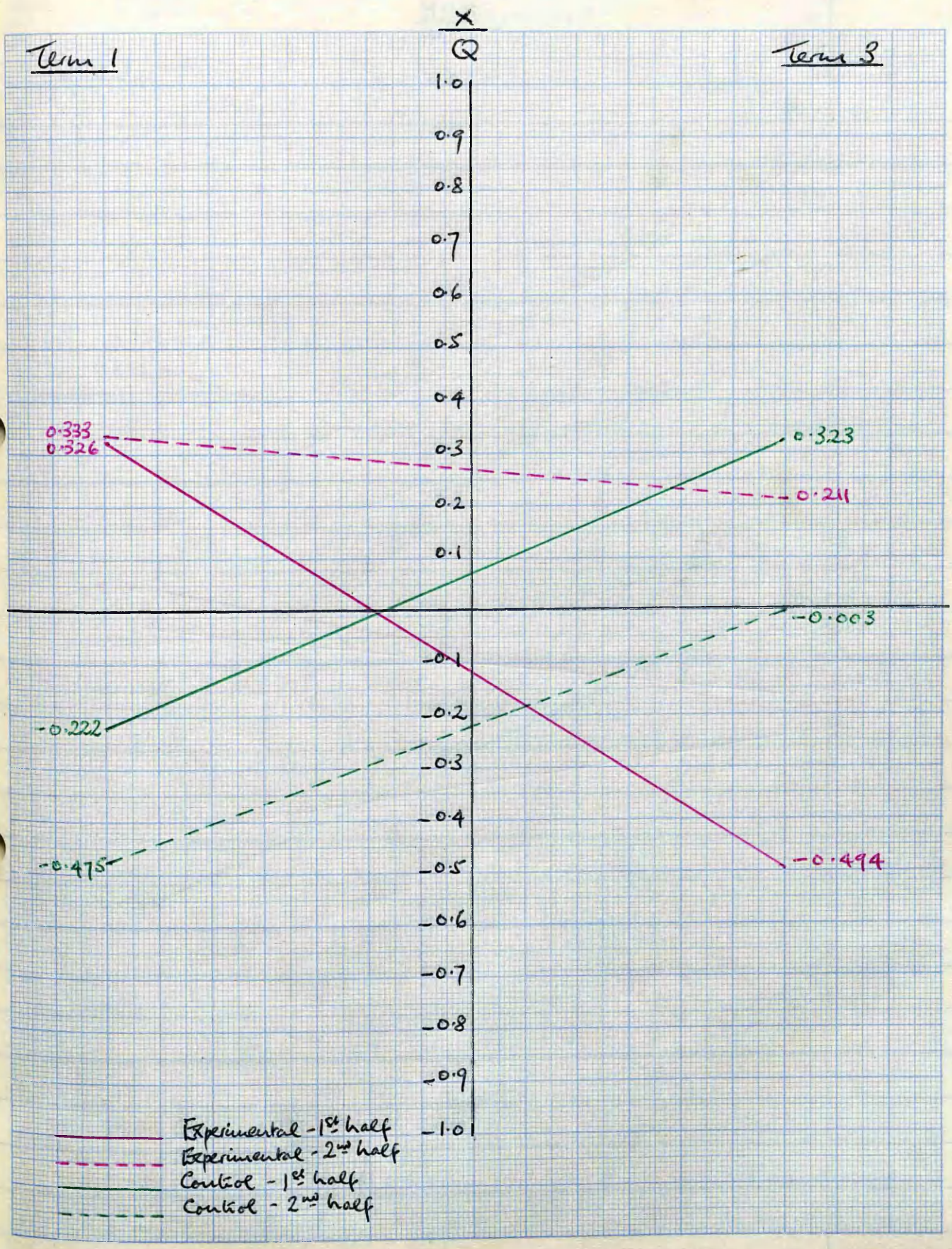
Q

Term 1

Term 3



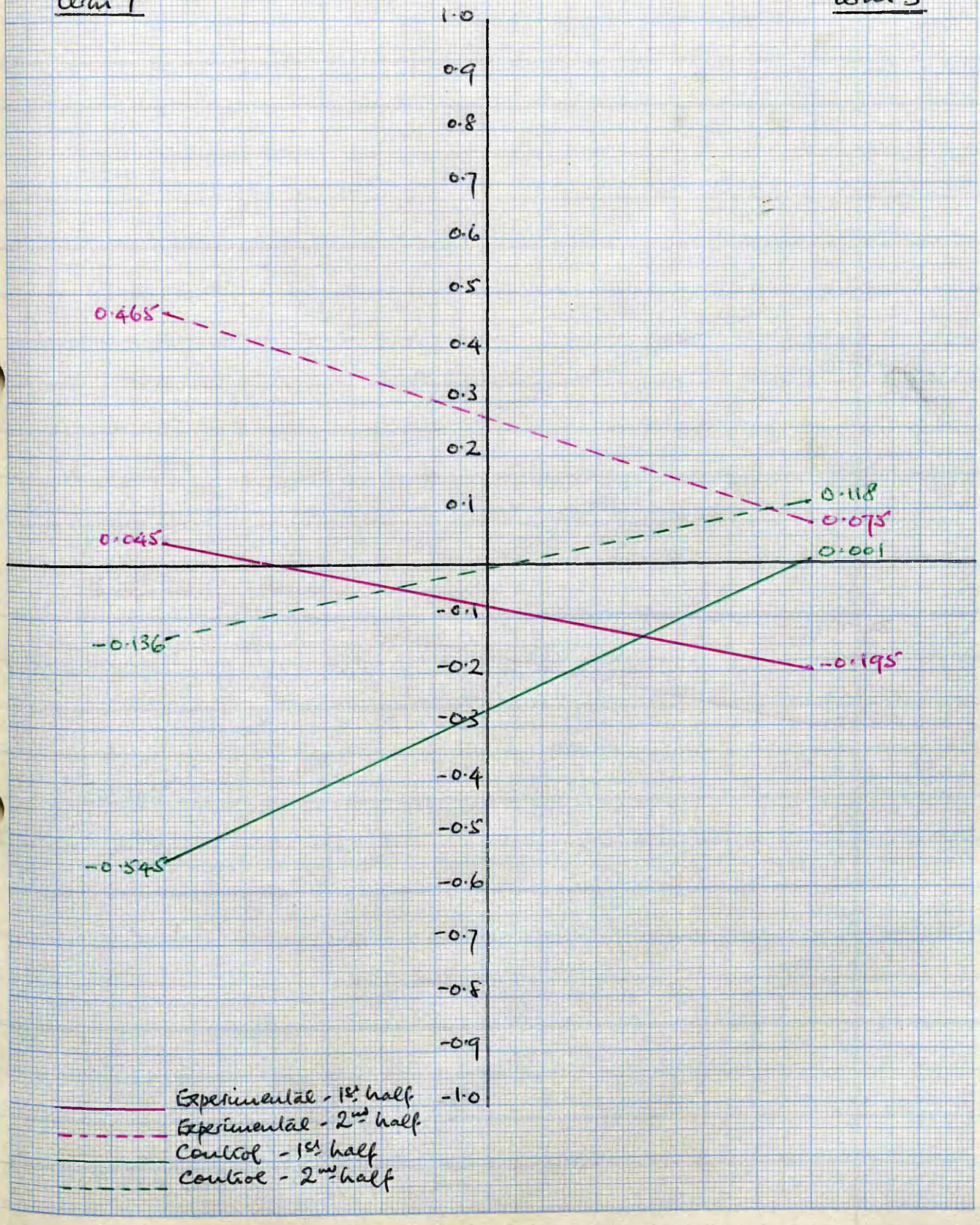

y
Q
Term 1
Term 3
1·0
0·9
0·8
0·7
0·6
0·5
0·4
0·3
0·2
0·1
0·825
0·147
-0·034
-0·060
-0·044
-0·1
-0·167
-0·2
-0·243
-0·3
-0·423
-0·4
-0·5
-0·6
-0·7
-0·8
-0·9
-1·0
Experimental - 1st half
Experimental - 2nd half
Control - 1st half
Control - 2nd half


X
Q
Term 1
Term 3
1·0
0·9
0·8
0·7
0·6
0·5
0·4
0·3
0·2
0·1
-0·1
-0·2
-0·3
-0·4
-0·5
-0·6
-0·7
-0·8
-0·9
-1·0
0·333
0·326
0·323
0·211
-0·003
-0·222
-0·475
-0·494
Experimental - 1st half
Experimental - 2nd half
Control - 1st half
Control - 2nd half

HOX / HOQ

Term 1

Term 3



n

Q incl n

Term 1

Term 3

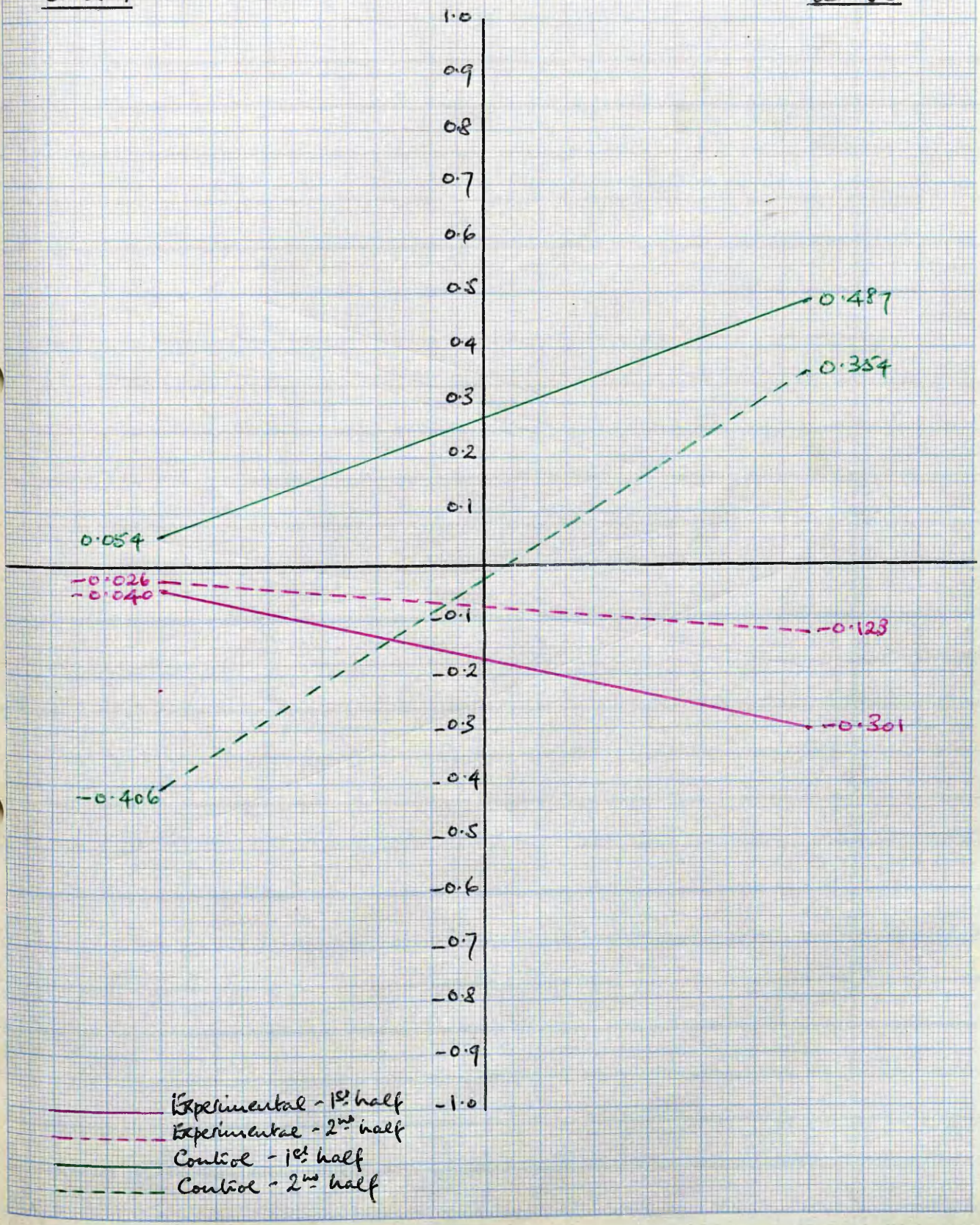


Table XVII

Group means of standardized scores: term 1 - term 3

(v) Analysis of variance based on term 3 scores.

In the final term of the research project, the deviation of any individual's score from the mean score of the population, on each of the selected skills, could feasibly have been influenced by the following factors:

- the treatment given (experimental v. control)
- the ranking (based on initial ability in the selected skills) of the pair to which he belonged
- the half term in which he was assigned to teaching practice

The design of the research experiment provided for a two-factor analysis of these sources of variance, with the influence of ranking examinable only within each separate half term (i.e. the pair effect was nested under the half term effect). WINER B J (Statistical Principles in Experimental Design, p. 303) sets out a model for an analysis of this kind, appropriate to a two-factor experiment with repeated measures on one factor. In concrete terms, the design of the Jordanhill experiment involved, not repeated measures on individuals, but single measures on each member of a pair of subjects. Winer's model was nevertheless considered to be relevant, since between-treatment differences within matched pairs may be taken as formally equivalent to between-tretament differences within individuals. In the Jordanhill design, individuals were matched and then randomly allocated within pairs to experimental or control treatments. From a statistical point of view, this procedure is identical with the situation described by Winer, in which each individual was exposed to successive treatments.

Similarly/

Similarly, nesting within levels of the second factor was characteristic of both designs; applicable to individuals in Winer's model and to matched pairs in the Jordanhill experiment.

The linear model for the analysis, as given by Winer, is

$$X = \mu + \alpha + \pi + \beta + \alpha\beta + \beta\pi + \varepsilon$$

where X = the individual's score

μ = mean score of the population

α = effect of the half term

π = pair effect within a half term

β = effect of treatment

$\alpha\beta$ = interaction between half term effect and effect of treatment

$\beta\pi$ = interaction between treatment effect and pair within half term effect

ε = random variation among individuals

For each skills category, the data for analysis were set out in the following form:

Table XVIII

Data for analysis of variance

Half terms (P)	Pair	Treatments (q)		Totals P
		Exptl. (B_1)	Control (B_2)	
1st Half term (A_1)	1	X1	X19	P1
	2	X2		P2
	3	X3		P3
	4			
	(n) 5			
	6			
	7			
	8			
	9	X9	X27	P9
2nd Half term (A_2)	10	X10	X28	P10
	11			
	12			
	13			
	(n) 14			
	15			
	16			
	17			
	18	X18	X36	P18
				G

Summary Totals

	Exptl. (B_1)	Control (B_2)	Totals
1st Half (A_1)	ΣX1-9	ΣX19-27	
2nd Half (A_2)	ΣX10-18	ΣX28-36	
Totals			

where A_1 and A_2 = first and second half term groups

p = number of levels of A

B_1 and B_2 = treatments (experimental and control)

q = number of levels of B

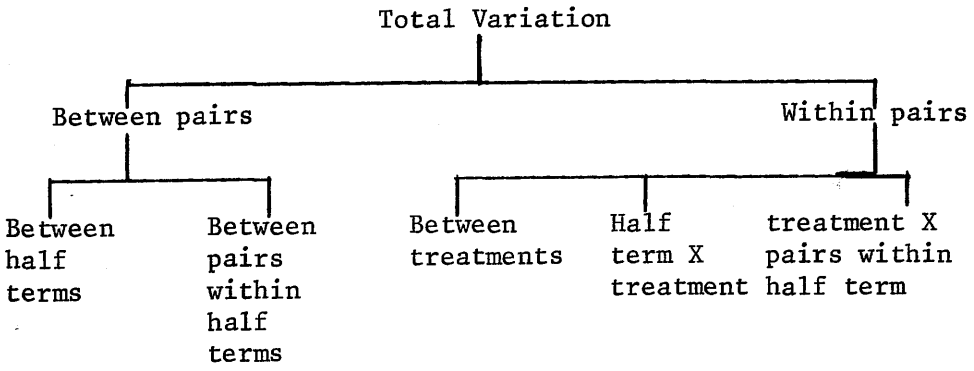
n = number of pairs in each half term group

X = individual scores

$G = \Sigma X$

P = total of scores for each pair

The breakdown of variation in the experiment can be shown as follows:



where Half term X treatment interaction = a measure of the extent to which the effect of half term placing (the level of A) varies according to the treatment received (the level of B)

Treatment X pairs within half term interaction = a measure of the extent to which, within a half term, the treatment effect varies according to the ranking of any particular pair.

The/

The form of the analysis of variance is summarized in Table XIX.

Table XIX

Summary of analysis of variance

Source of variation	Degrees of freedom		Expected Mean Square
Between pairs	np - 1	17	$\sigma^2_\varepsilon + q\sigma^2_\pi + nq\sigma^2_\alpha$
A (between half terms)	p - 1	1	
Pairs within half terms	p (n - 1)	16	$\sigma^2_\varepsilon + q\sigma^2_\pi$
Within pairs	np (q - 1)	18	$\sigma^2_\varepsilon + \sigma^2_{\beta\pi} + np\sigma^2_\beta$
B (between treatments)	q - 1	1	
AB (half term x treatment)	(p - 1) (q - 1)	1	$\sigma^2_\varepsilon + \sigma^2_{\beta\pi} + n\sigma^2_{\alpha\beta}$
B x pairs within half terms (residual error)	p (n - 1) (q - 1)	16	$\sigma^2_\varepsilon + \sigma^2_{\beta\pi}$

The computational symbols used for the analysis were

(1) $\frac{G^2}{npq}$ (4) $\frac{\sum B^2}{np}$

(2) $\sum X^2$ (5) $\frac{\sum (AB)^2}{n}$

(3) $\frac{\sum A^2}{nq}$ (6) $\frac{\sum P^2}{q}$

(n.b. (2) - (1) = total sum of squares)

The three variance ratios to be tested for significance were

Sources of variance between pairs

$$\frac{\text{Between half terms}}{\text{Between pairs within half terms}}$$

Sources of variance within pairs

$$\frac{\text{Between treatments}}{\text{Treatment X pairs within half terms}}$$

$$\frac{\text{Between (half terms X treatments)}}{\text{Treatments X pairs within half terms}}$$

The/

The application of the analysis of variance to each of the skills categories is set out in Table XX (a-h)

Table XX(a)

Analysis of variance: $\frac{HOQ}{Q}$

Source of variation	Computational formula	SS	df	MS	F
Between pairs	(6)-(1)	= 2949.56	17		
Between half terms (A)	(3)-(1)	= 0.12	1	0.12	0.00065
Pairs within half terms	(6)-(3)	= 2949.44	16	184.34	
Within pairs	(2)-(6)	= 5971.73	18		
Between treatments (B)	(4)-(1)	= 3164.06	1	3164.06	21.66*
A B	(5)-(3)-(4)+(1)	= 470.17	1	470.17	3.22
Treatments X pairs within half terms	(2)-(6)-(5)+(3)	= 2337.5	16	146.09	

*Significant at .001 level

Table XX(b)

Analysis of variance: $\frac{LOQ}{Q}$

Source of variation	Computational formula	SS	df	MS	F
Between pairs	(6)-(1)	= 6053.31	17		
Between half terms (A)	(3)-(1)	= 13.82	1	13.82	0.037
Pairs within half terms	(6)-(3)	= 6039.49	16	377.47	
Within pairs	(2)-(6)	= 11122.97	18		
Between treatments (B)	(4)-(1)	= 6133.5	1	6133.5	20.11*
A B	(5)-(3)-(4)+(1)	= 109.9	1	109.9	0.36
Treatments X pairs within half terms	(2)-(6)-(5)+(3)	= 4879.57	16	304.97	

*Significant at .001 level

Table XX(c)

Analysis of variance: XP / X

Source of variation	Computational formula	SS	df	MS	F
Between pairs	(6)-(1)	= 15655.02	17		
Between half terms (A)	(3)-(1)	= 5198.41	1	5198.41	7.95*
Pairs within half terms	(6)-(3)	= 10456.61	16	653.54	
Within pairs	(2)-(6)	= 13894.06	18		
Between treatments (B)	(4)-(1)	= 1369.0	1	1369.0	1.9
A B	(5)-(3)-(4)+(1)	= 971.36	1	971.36	1.35
Treatments X pairs within half terms	(2)-(6)-(5)+(3)	= 11553.7	16	722.11	

* significant at .02 level

Table XX(d)

Analysis of variance: Pr / Q

Source of variation	Computational formula	SS	df	MS	F
Between pairs	(6)-(1)	= 406.09	17		
Between half terms (A)	(3)-(1)	= 10.14	1	10.14	0.41
Pairs within half terms	(6)-(3)	= 395.95	16	24.75	
Within pairs	(2)-(6)	= 1197.08	18		
Between treatments (B)	(4)-(1)	226.51	1	226.51	3.82
A B	(5)-(3)-(4)+(1)	= 22.55	1	22.55	0.38
Treatments X pairs within half terms	(2)-(6)-(5)+(3)	= 948.02	16	59.25	

Table XX(g)

Analysis of variance: HOX / HOQ

Source of variation	Computational formula	SS	df	MS	F
Between pairs	(6)-(1)	= 6879.76	17		
Between half terms (A)	(3)-(1)	= 145.77	1	145.77	0.35
Pairs within half terms	(6)-(3)	= 6733.99	16	420.87	
Within pairs	(2)-(6)	= 6711.73	18		
Between treatments (B)	(4)-(1)	= 55.66	1	55.66	0.13
A B	(5)-(3)-(4)+(1)	= 22.65	1	22.65	0.05
Treatments X pairs within half terms	(2)-(6)-(5)+(3)	= 6633.42	16	414.59	

Table XX(h)

Analysis of variance: n / Q incl n

Source of variation	Computational formula	SS	df	MS	F
Between pairs	(6)-(1)	= 1490.21	17		
Between half terms (A)	(3)-(1)	= 0.34	1	0.34	0.0037
Pairs within half terms	(6)-(3)	= 1489.87	16	93.12	
Within pairs	(2)-(6)	= 1402.28	18		
Between treatments (B)	(4)-(1)	= 268.41	1	268.41	3.84
A B	(5)-(3)-(4)+(1)	= 16.14	1	16.14	0.23
Treatments X pairs within half terms	(2)-(6)-(5)+(3)	= 1117.73	16	69.86	

Table XX(e)

Analysis of variance: $\frac{Y}{Q}$

Source of variation	Computational formula	SS	df	MS	F
Between pairs	(6)-(1)	= 2229.62	17		
Between half terms (A)	(3)-(1)	= 331.85	1	331.85	2.8
Pairs within half terms	(6)-(3)	= 1897.77	16	118.61	
Within pairs	(2)-(6)	= 3034.43	18		
Between treatments (B)	(4)-(1)	= 208.81	1	208.81	1.25
A B	(5)-(3)-(4)+(1)	= 146.81	1	146.81	0.88
Treatments X pairs within half terms	(2)-(6)-(5)+(3)	= 2678.81	16	167.43	

Table XX(f)

Analysis of variance: $\frac{X}{Q}$

Source of variation	Computational formula	SS	df	MS	F
Between pairs	(6)-(1)	= 1024.43	17		
Between half terms (A)	(3)-(1)	= 23.04	1	23.04	0.37
Pairs within half terms	(6)-(3)	= 1001.39	16	62.59	
Within pairs	(2)-(6)	= 1653.03	18		
Between treatments (B)	(4)-(1)	= 57.76	1	57.76	0.65
A B	(5)-(3)-(4)+(1)	= 169.87	1	169.87	1.91
Treatments X pairs within half terms	(2)-(6)-(5)+(3)	= 1425.4	16	89.09	

(vi) Probability related to experimental v. control comparisons of superiority.

Control and experimental group performances (as measured by t-test scores) on each of the eight criteria in terms 1 and 3 were compared with the following results:

Table XXI

Experimental v. Control comparisons of superiority: term 1, first and second halves

Criterion	Superiority 1st Half		Superiority 2nd Half	
$\frac{HOQ}{Q}$	E			C
$\frac{LOQ}{Q}$	E		E	
$\frac{XP}{X}$		C	E	
$\frac{Pr}{Q}$	E		E	
$\frac{Y}{Q}$	E		E	
$\frac{X}{Q}$		C		C
$\frac{HO\ X}{HOQ}$	not scored			C
$\frac{n}{Q \text{ incl. } n}$	E			C
Totals	5	2	4	4

E = experimental
C = control

Table XXII

Experimental v. Control comparisons of superiority: term 3, first and second halves

Criterion	Superiority 1st Half		Superiority 2nd Half	
$\frac{HOQ}{Q}$	E		E	
$\frac{LOQ}{Q}$	E		E	
$\frac{XP}{X}$	E		E	
$\frac{Pr}{Q}$	E		E	
$\frac{Y}{Q}$	E		E	
$\frac{X}{Q}$	E			C
$\frac{HO\ X}{HOQ}$	E		E	
$\frac{n}{Q \text{ incl. } n}$	E		E	
Totals	8	-	7	1

Table XXIII

Experimental v. control comparisons of superiority: gain scores, term 1-term 3, first and second halves

Criterion	Superiority 1st Half		Superiority 2nd Half	
$\frac{HOQ}{Q}$	E		E	
$\frac{LOQ}{Q}$	E		E	
$\frac{XP}{X}$	E		E	
$\frac{Pr}{Q}$	E			C
$\frac{Y}{Q}$		C	E	
$\frac{X}{Q}$	E		E	
$\frac{HO\ X}{HOQ}$	E		E	
$\frac{n}{Q \text{ incl. } n}$	E		E	
Totals	7	1	7	1

The distribution of probability related to the possible levels of superiority as between the experimental and control groups may be expressed as binominal expansions:

- for Term 1, first half: $(1+1)^7 = 1+7+21+35+35+21+7+1 = 128$

- for all other occasions: $(1+1)^8 = 1+8+28+56+70+56+28+8+1 = 256$

Thus the total probability factor (p) associated with the levels of superiority achieved by the experimental group in terms 1 and 3 and in term 1-term 3 gain scores may be shown in the form set out in Table XXIV. These values of p were examined for significance/

significance at the .05 and .01 levels. 2-tailed tests were applied to the term 1 scores and 1-tailed tests to the term 3 scores and term 1-term 3 gains, on the assumption that the operation of the independent variable would have favoured the experimental group.

Table XXIV

Values of p associated with experimental group levels of superiority in terms 1 and 3

Occasion	Value of p	Level of significance
Term 1		
1st half	$p = \frac{29}{128} \times 2 = 0.454$	N.S.
2nd half	$p = \frac{128}{256} \times 2 = 1.0$	N.S.
Term 3		
1st half	$p = \frac{1}{256} = 0.004$	$<.01$
2nd half	$p = \frac{9}{256} = 0.035$	$<.05$
Term 1-term 3		
1st half	$p = \frac{9}{256} = 0.035$	$<.05$
2nd half	$p \frac{9}{256} = 0.035$	$<.05$

(vii) Percentage of teacher talk in term 1 and term 3 recorded lessons.

The percentage of teacher talk was computed for the ten-minute discussion sections of the control and experimental group recorded lessons in terms 1 and 3. Results were as follows:

Table XXV

Experimental and control group percentages of teacher talk: term 1

Exp. Group	1st Half	2nd Half	Cont. Group	1st Half	2nd Half
	%	%		%	%
1	67.0	80.8	1	76.7	68.4
2	79.8	84.4	2	77.7	72.0
3	81.6	81.0	3	80.0	58.2
4	54.0	65.6	4	60.2	77.0
5	73.0	69.9	5	81.0	78.4
6	65.0	84.6	6	81.7	82.8
7	85.7	56.5	7	93.0	81.2
8	61.5	75.8	8	77.7	65.9
9	73.4	68.6	9	74.2	87.7
Averages	71.22	74.13	Averages	78.02	74.62

Table XXVI

Experimental and control group percentages of teacher talk: term 3

Exp. Group	1st Half	2nd Half	Cont. Group	1st Half	2nd Half
	%	%		%	%
1	64.7	74.5	1	83.1	73.4
2	74.5	81.1	2	70.4	78.6
3	61.5	62.3	3	85.5	81.7
4	69.5	86.0	4	64.5	71.3
5	75.7	69.8	5	82.2	67.0
6	57.0	69.4	6	71.7	73.2
7	72.1	73.0	7	87.1	81.5
8	46.0	74.3	8	78.8	56.9
9	62.5	73.4	9	80.3	86.0
Averages	64.83	73.76	Averages	78.18	74.4

The significance of differences between mean performances was examined by the application of t-tests, using the formula for matched subject designs. Results are set out in Tables XXVII and XXVIII.

Table XXVII

Term 3-term 1 t-values for percentages of teacher talk: experimental and control groups

	Experimental	Sig.	Control	Sig.	
		P<		P<	
1st Half	1.782		-0.074		+favours term 3
2nd Half	0.087		0.061		

Table XXVIII

Experimental-control t-values for percentage of teacher talk: terms 1 and 3

	Exp.-Cont. Term 1	Sig.	Exp.-Cont. Term 3	Sig.	
		P<		P<	—favours experimental group
1st Half	-2.950	.02[x]	-3.229	.01[+]	
2nd Half	-0.090		-0.161		

[x] 2-tailed [+] 1-tailed

B./

B. Correlations between term 1 codings and tutors' term 1 assessments.

In each half of term 1, students had been ranked for initial ability in the selected questioning skills on the basis of the sums of scores on those skills awarded them by tutors. This ranking was compared with their overall ranking based on their coded performance on related skills in the recorded term 1 lessons. This latter overall ranking was arrived at by taking, for each student, the sum of rankings on the six ratios of $\frac{HO\ Q}{Q}$, $\frac{LO\ Q}{Q}$, $\frac{xP}{X}$, $\frac{Pr}{P}$, $\frac{Y}{Q}$ and $\frac{n}{Q \text{ incl. } n}$

(n.b. Rank order was directly related to percentage scores for $\frac{HO\ Q}{Q}$, $\frac{xP}{X}$ and $\frac{Pr}{Q}$; and inversely related for $\frac{LO\ Q}{Q}$, $\frac{Y}{Q}$ and $\frac{n}{Q \text{ incl. } n}$.)

Correlations between tutor and coder overall rankings were determined, using the formula

$$\rho = 1 - \frac{6\Sigma d^2}{n(n^2-1)}$$

where d = difference between rankings given to each student

n = number of students in each group

Table XXIX gives the rank order correlations for each half term group.

Table XXIX

Rank order correlations of term 1 performances on selected questioning skills (tutors' and coder's assessments)

	Experimental	Control
1st Half	-.04	.22
2nd Half	-.54	-.1

In view of these very low correlations, additional rank order correlations, related to selected skills performance, were determined as follows:

(i)/

(i) between experimental and control group term 1 rankings, based on codings of recorded lessons.

(N.B. In terms of tutor ratings, experimental and control groups exhibited perfect positive correlation in term 1.)

(ii) for each group of students, between rankings based on assessments made by tutors on their first and second visits.

(N.B. Initial matching of students had been based on the average of three tutor assessments.)

The formula used for these correlations, embodying corrections for tied rankings, was

$$\rho = \frac{\Sigma' x^2 + \Sigma' y^2 - d^2}{\sqrt[2]{\Sigma x^2 \times \Sigma y^2}}$$

Details of these two sets of rank order correlations are set out in Tables XXX and XXXI.

Table XXX

Experimental v control group correlations based on codings of recorded term 1 lessons

	Experimental v Control
1st Half term 1	.45
2nd Half term 1	.34

Table XXXI

Rank order correlations based on first and second assessments by tutors in term 1

	Experimental group	Control group
1st Half	.15	.56
2nd Half	.55	.42

C. Measurements based on tutors' assessments.

(i) The Appraisal Guide (see Appendix C) used by tutors on their visits to students in schools provided both for broad assessments of overall teaching competence and for more detailed assessments of performance on each of the selected questioning skills. t-tests were applied to examine the significance of any differences between experimental and control groups which emerged from these tutor ratings. Once again, the formula for matched subject designswas used. Results are set out in Tables XXXII and XXXIII.

Table XXXII

t-values derived from tutors' overall lesson rating of experimental and control groups, term 1 and term 3

	Experimental - Control			
	Term 1	Sig.	Term 3	Sig.
		P<		P<
1st Half	0.437		0.930	
2nd Half	-1.259		-1.750	

N.B. All tests of significance were 2-tailed.

Table XXXIII

t values derived from tutors' term 3 ratings of performance on questioning skills

SKILLS	Experimental - Control			
	1st Half	Sig.	2nd Half	Sig.
		P<		P<
One-word answers	-0.700		-0.760	
Yes/no questions	-1.871	.05	-1.000	
Optimum use of HOQ	0.532		-0.667	
Allowing time for response	-0.817		-0.665	
Prompting	-0.158		-1.206	
Probing	-0.521		-1.306	
Redirecting	-0.853		-0.935	

N.B. All tests of significance were 1-tailed

It should be noted that the seven skills rated by tutors are not identical with the behaviours selected for quantitative measurement in the coding procedure, but are closely related to them. Thus

- avoiding over-use of questions relates to low $\frac{LOQ}{Q}$ coefficient demanding one-word answers

- avoiding over-use of questions relates to low $\frac{Y}{Q}$ coefficient demanding yes/no response

- optimum use of HOQ relates to high $\frac{HOQ}{Q}$ coefficient

- allowing/

- allowing time for response relates to low $\frac{n}{Q \text{ incl. } n}$ coefficient

- prompting relates to high $\frac{XP}{X}$ coefficient

- probing relates to high $\frac{Pr}{Q}$ coefficient

Coding of Redirection was not processed, due to the infrequent use of this skill. The proportion of questions not attracting a response $\left(\frac{X}{Q}\right)$ is not directly related to any tutor assessment.

(ii) Since the general pattern of tutor term 3 ratings ran counter to the pattern of evidence derived from the coding of term 3 lessons, the History department proceeded with the plan to make their own ratings of the questioning skills exhibited in the previously coded discussion section of these recordings (see chapter 4, section (vii), (2)). Tutors were paired on the basis of practical convenience in sharing tape recorders. No attempt was made to promote or avoid the pairing of tutors likely to agree in their assessments. Four of the five pairs of tutors had seven lessons to assess. (3-4 experimental, 3-4 control), while the remaining pair had eight (4 experimental, 4 control). Within these constraints, lessons were randomly distributed. Tutors undertook to work independently in making their assessments and to preserve confidentiality throughout the experiment. Moreover, they remained unaware of which students were in the control and which in the experimental group.

A/

A rating form was devised (see Appendix P) which brought the criteria as closely as possible into line with the criteria used for the quantitative measurements made by the researcher. The definitions of the skills included in the rating form notes were taken from the Ground Rules for coding. The aim was that tutors should be guided towards measuring the same behaviours as those defined by the Ground Rules, while taking account of the concept of "appropriateness".

Product-moment correlations were carried out in respect of each criterion, using the formula

$$r = \frac{\Sigma' xy}{\sqrt{\Sigma' x^2 \times \Sigma' y^2}}$$

where x = assessment by first tutor in pair

y = assessment by second tutor in pair

$$\Sigma' xy = \Sigma xy - \frac{(\Sigma x)\ (\Sigma y)}{n}$$

$$\Sigma' x^2 = \Sigma x^2 - \frac{(\Sigma x)^2}{n}$$

$$\Sigma' y^2 = \Sigma y^2 - \frac{(\Sigma y)^2}{n}$$

n = number of lessons assessed

These correlations were subsequently extended to compare the ratings of each of the tutors with the coder's measurement of the related criterion in the same recorded lesson. Thus for each group of seven or eight lesson extracts there were three comparisons applied to each criterion:

r ab - first tutor in pair: second tutor in pair

r ac - first tutor in pair: coder

r bc - second tutor in pair: coder

The results of these product-moment correlations are set out in Table XXXIV.

Table XXXIV

Product-moment correlations of tutor pair and tutor-coder ratings of questioning skills in term 3 recorded lessons

Criterion	Comparison	Lesson groups				
		1(n=8)	2(n=7)	3(n=7)	4(n=7)	5(n=7)
Optimum HOQ	r ab	.702	.048	.616	-.227	1.0
	r ac	.231	.788	-.286	.854	.462
	r bc	.167	.342	-.196	-.374	.462
Restricted LOQ	r ab	.317	-.025	.240	.445	.943
	r ac	.378	.117	.172	.352	.555
	r bc	.806	.356	.501	-.527	.607
Pausing	r ab	.346	-.129	.242	-.146	.923
	r ac	.334	.060	.649	.038	.794
	r bc	.640	-.170	.839	-.194	.850
Prompting	r ab	.455	.283	.391	-.287	.907
	r ac	.025	.066	.342	-.322	.102
	r bc	-.461	.733	.674	-.429	.188
Probing	r ab	.548	-.291	.663	.232	.950
	r ac	-.235	.505	.489	.530	.075
	r bc	.483	-.324	.799	.472	.167
Re-directing	r ab	.587	-.274	.481	.359	.932
	r ac, r bc	Coder's measurements not processed				

In view of the marked inconsistency of the tutor pair correlations, it was decided to proceed no further with the measurements listed in Chapter 4, page 116. However, the performances of control and experimental groups, as assessed by the mean of tutor pair ratings of/

of the recorded lessons, were compared, and the differences were expressed in terms of the number of skills on which each group showed superiority. The following results were obtained:

Table XXXV

Experimental v. control comparisons of superiority based on tutors' assessments of term 3 recorded lessons

Criterion	Superiority		Superiority	
	1st Half term		2nd Half term	
Optimum HOQ	E		E	
Redirected LOQ	E		E	
Pausing	E		E	
Prompting	E		E	
Probing	E		E	
Redirecting	E			C
Totals	6	-	5	1

E = Experimental

C = Control

D. Analysis of student and tutor reactions to microteaching.

In view of the small population of students and tutors involved in the project, no statistical procedures were applied to the questionnaire responses. Information is summarized under the two main headings of

- questions demanding ticked responses
- questions requiring comment

The three questionnaires are reproduced in full as Appendices J, K and L.

(i) Term 2 questionnaire (students)

Table XXXVI

Term 2 questionnaire responses from first and second half term experimental groups

Ref.	Question	Response		No response	
		1st	2nd	1st	2nd
A1	Did you find time to read handout				
a	- before or during term 1 practice				
	yes	8	6		
	no	1	3		
b	- before term 2 microteaching				
	yes	8	5		
	no		3	1	1
2	Analysis of questioning skills was				
	- obscure and of little help				
	- clearly set out and of practical use	4	4		
	- clearly set out, but too theoretical	5	3		
	- inadequately explained, but apparently relevant		1		1
3	Did you refer back to handout				
	- occasionally	6	3		
	- not at all	3	5		1
4	Did you refer to handout before assessment lesson				
	- yes	4	2		
	- no	5	6		1
B1	Reaction to videotape models				
	- made nature of skills clearer	7	7		
	- added nothing to understanding	2	1		
	- left me more confused		1		
2	Reaction to videotape model commentary				
	- essential to understanding		2		
	- acceptable as a reinforcement	8	7		
	- took up undue time	1			

Table XXXVI (Contd.)

Ref.	Question	Response		No response	
		1st	2nd	1st	2nd
B3	Time spent on viewing models				
	- about right	6	5		
	- would have liked longer	3	4		
4	Viewing of models immediately before microteaching practice				
	- to be preferred	3	3		
	- gap between viewing and practice preferable	5	6	1	
5	Videotape model component was				
	- most useful of all				
	- very useful	2			
	- reasonably useful	7	8		
	- not very useful		1		
	- no use at all				

Table XXXVI (Contd.)

Ref.	Question	Response		No response	
		1st	2nd	1st	2nd
C1 a	Briefing letter was				
	- all that was needed for planning	8	3	1	
	- more detail useful		6		
b	Time spent planning microteach lessons				
	- Lesson A				
	- less than ½ hour	1	2		
	- ½ hour - 1 hour	3	4		
	- 1 hour - 2 hours	3	2		
	- more than 2 hours	2	1		
	- Lesson B				
	- less than ½ hour		2		
	- ½ hour - 1 hour	4	2		
	- 1 hour - 2 hours	4	4	1	
	- more than 2 hours		1		
c	Planning of microteach lesson was				
	- relatively easy	3	4		
	- rather difficult	6	5		
d	Difficulty of c due to				
	- lack of guidance over subject matter	2	2		
	- 7-8 minute presentation time	4	5		
	- concentrating on specific skills	3	4		
2 a	Reactions to microteaching				
	- artificial (few pupils)	3	5		
	- concentration on skills helped by lack of disciplinary, organisational problems	5	6		
	- 7-8 minutes was adequate	3	2		
	- teaching period too short	1	5		
	- felt unspontaneous and self-conscious	3	3		
	- concentrating on few skills a help	8	7		
	- microteaching pleasant but irrelevant	2	1		
	- microteaching a useful preparation for class teaching	5	5		
	- distracted by microphones, etc.		1		
	- no clear idea of skills				

Table XXXVI (Contd.)

Ref.	Question	Response		No response	
		1st	2nd	1st	2nd
C2 c	Concentration on two complex skills				
	- able to cope	6	5		
	- preferred one skill only per lesson	3	3		
3 a	Reactions to group review of lessons				
	- peer comments helpful	3	6		
	- listening to peer teaching useful	7	7		
	- prefer review with tutor alone				
	- no strong feelings	3	2		
b	Reactions to tutor's presence				
	- presence of tutor essential	6	5		
	- teaching alone preferred, but tutor welcome at review	2	1		
	- tutor presence an embarrassment	2			
	- presence or absence of tutor not significant	1	4		
c	Reactions to tutor's role				
	- too authoritarian				
	- more positive guidance needed	1			
	- comments very helpful	7	7	1	2
d	Completion of replay evaluation forms				
	- helped me to concentrate on skill	4	5		
	- a distraction to concentration	5	3		
e	Commenting on one's own performance				
	- embarrassing ordeal				
	- difficult to be objective	7	4		
	- chance of self-criticism welcomed	2	5		
	- better to spend time listening to others' comments	2	3		
f	Review sessions were				
	- valuable and right length	3	5		
	- valuable, should have been longer	4			
	- not very valuable	2			
	- valuable, could have been shortened		4		

Table XXXVI (Contd.)

Ref.	Question	Response		No response	
		1st	2nd	1st	2nd
C4 a	Replanning between teach and reteach				
	- major changes made	2			
	- a few changes made	6	6		
	- kept to same plan	1	3		
b	Reactions to reteach lesson				
	- helped to improve confidence	5	3		
	- added nothing of value				
	- correction of initial weaknesses facilitated	6	5		
	- had little effect on use of skill	2	2		

General comments from term 2 questionnaire:

A5 - Disagreement with handout statements

1st Half - Nil response

2nd Half - Nil response

A6 - Comments on handout

1st Half - Seemed too theoretical at first, but became clearer in light of experience.

2nd Half - Skills could have been more "positively" presented.

- Distinction between Prompting and Probing not too clear.

B6 - Comments on videotape models

1st Half - Commentary links helped in analysis of teaching sequences.

- Commentary useful, but extended examples would have helped.

- It would have been better to view models one section at a time, before practising each skill. As it was, student tended to confuse Prompts and Probes.

- It would have been useful at end of second day's microteaching, to supplement models by showing experienced teacher using skills in complete lesson with normal sized class.

2nd Half - "Could be irritating - always model classes, model lessons - all in stark contrast to the reality in schools".

- Models should have shown a 7-8 minute microteaching lesson, not extracts from a "normal" lesson.

- Models appeared to be set in "ideal" condition. Doubtful about success of techniques with pupils of low ability.

- "Was the model too good?"

- Models seemed artificial. Skills appeared easier to use than in fact they were.

C2b - Reactions to experience of microteaching

1st Half - Microteaching did not tackle problem of retaining interest of less able pupils while one or two were answering questions.

- Problem of controlling discussion while not inhibiting pupils should be tackled.

- Pupils seemed inhibited by microteaching. A teacher should know his pupils (impossible in microteaching) and adapt his question technique accordingly.

2nd Half - Pupils appeared above average - this may have reduced usefulness of experience.

C3a - Reactions to group review sessions

1st Half - Embarrassment passed quickly.

- Visual feedback would have helped understanding of pupils' reactions better than audio.

- Readier criticism by students would have made sessions more stimulating.

2nd Half - Nil response.

C3b - Reactions to tutors' presence

1st Half - Tutor could analyse lesson in relation to desired criteria more clearly than students. Therefore his presence was invaluable.

- More detailed comments from tutor on teach lesson would have helped planning of re-teach.

- If a lesson had gone badly, tutor's presence might have been an embarrassment.

2nd Half - During student teaching, tutor (and technician) should have been concealed. Student would have felt more "in control".

C3c - Comments on tutor's role

1st Half - "Instant judgments" disliked. Preferable to delay review for a few days.

- Replay/

- Replay of recording alone would have been useful, but tutor could "put his finger more clearly on particular points".

2nd Half - Tutor constructive and not threatening.

- Playback of specific sections of tape provided useful pointers.

- Tutor from student's own subject department may prove inhibiting to sensitive students.

- Improvement between teach and re-teach depended on tutor's comments.

- Criticisms of historical inaccuracies in content irrelevant to this exercise.

D - General Comments

1st Half - Importance of selecting appropriate material should be stressed more strongly.

- Handout and model tapes would be useful in term 1, before first teaching practice.

- Microteaching sessions helped
  - to clarify nature of questioning skills
  - to slow down my teaching pace
  - to give me confidence to pause for probing and re-directing. "In the past, I had been too anxious to provide the answers myself and afraid lest probing would lead nowhere. I would hope that this aspect of the micro-teaching experience could be consciously retained by me in any ordinary classroom lesson."

- Playbacks helped to achieve awareness of faults and to relate tutor criticism to actual examples.

2nd Half - Microteaching lessons should be kept short, otherwise review sessions become tedious.

- Unsure whether, by end of day, use of skills was achieved or their nature understood.

- Day seemed very long. It would have been better split into two.

- "I wonder how many departmental heads will be convinced of the merit of depth of understanding achieved by practice of the probing and re-directive techniques, if subject area targets are not achieved. The normal enquiry of the student teacher, mainly with certificate classes, is 'How far did you get?' Not 'What depth of discussion and understanding did you attain?'"

(ii) Term 3 questionnaire (students)

Table XXXVII

Term 3 questionnaire responses from first and second half term experimental groups

Ref.	Question	Response		No Response	
		1st	2nd	1st	2nd
A1	Conscious attempts to apply questioning skills in term 3 teaching practice				
	- not at all				
	- occasionally	4	3		
	- consistently	5	6		
2	Degree of difficulty or success in applying skills				
	- restricting one-word answer questions				
	- very difficult		1		
	- rather difficult	1	3		
	- moderate success	3	3		
	- good success	5	1		
	- restricting yes/no questions				
	- very difficult				
	- rather difficult	1			
	- moderate success	3	3		
	- good success	5	6		
	- optimum use of HOQ				
	- very difficult	1			
	- rather difficult	1	3		
	- moderate success	5	6		
	- good success	2			
	- pausing (for adequate response)				
	- very difficult				
	- rather difficult	7	4		
	- moderate success	2	4		
	- good success		1		
	- prompting				
	- very difficult				
	- rather difficult	1	1		
	- moderate success	6	8		
	- good success	1		1	
	- probing				
	- very difficult				
	- rather difficult	3	5		
	- moderate success	6	4		
	- good success				
	- re-directing				
	- very difficult	1	1		
	- rather difficult	2	3		
	- moderate success	6	5		
	- good success				

Table XXXVII (Contd.)

Ref.	Question	Response		No Response	
		1st	2nd	1st	2nd
A3	Reasons for difficulty in applying skills				
	- conflict with normal teaching style and personality		2		
	- question techniques failed to produce satisfactory pupil response level	1	2		
	- disciplinary problems, class size, etc. inhibited use of skills	6	3	2	2
4	Effect of skills on length and quality of pupil responses				
	- no noticeable difference				
	- slight improvement	4	4		
	- clearly marked improvement	5	4		1
5	Transfer of skills from microteaching to full-scale class lesson				
	- transfer impossible				
	- difficult, but possible	5	8		
	- no difficulty	4	1		
B1	Assessing value of microteaching as element in graduate training year				
	- of great value. More time should be be spent on it	5	4		
	- interesting, but no help with practical teaching problems	1	1		
	- reasonably valuable. Right amount of time spent on it	3	4		
	- of no value				
2	Retention of microteaching in future graduate course work				
	- recommended	9	9		
	- not recommended				
3	(If "recommended") should microteaching be applied				
	- initially, to all graduates	5	6		
	- reinforcement, for all graduates at midpoint/				

Table XXVII (Contd.)

Ref.	Question	Response		No Response	
		1st	2nd	1st	2nd
	midpoint of year	5	9		
	- remedially, for graduates weak in specific skills	8	4		
B4	Of the three main elements of micro-teaching: (a) definition and analysis of skills (b) practice of skills (c) self-assessment based on feedback which would you include in the graduate course?				
	- (a) without (b) or (c)		1		
	- (a) and (b), but not (c)	1			
	- (a), (b) and (c) as a complete cycle	8	8		

General comments from term 3 student questionnaire:

A3 - Problems associated with application of skills

1st Half - Tendency was to concentrate on brighter pupils (class as a whole became bored if too much time spent questioning weaker pupils).

- Teaching practice artificial. Skills easier to apply if you know the pupils.

- Disciplinary problems lead to concentration on recall questions for disciplinary reasons.

- "As far as I could see, many of the pupils were not used to 'higher order' questioning. They had got into the habit of answering yes/no type answers, because in some cases this was the only type of question their usual teachers asked."

- Temptation was to 'get through' the material and thus cut down time on questioning.

2nd Half - H.O.Q.'s difficult to use successfully with lower ability classes.

- "As regards pausing, I got the impression in some classes that the children were wondering if you had forgotten what was to come next."

- Re-direction difficult, because children loath to get involved in discussion.

- Difficult to combine use of questioning skills with task of getting through the material. "Often I found myself not getting enough work done relevant to the progress of the class."

- Only problem was the laziness, reticence, etc. - of some classes.

- 'One-word answer' questions can sometimes be effective in reaching a point which teacher wishes to emphasize.

- Many pupils are unused to being questioned. They need to be trained to take questions seriously.

B5 - Applications and modifications of microteaching in future graduate courses.

1st Half - Skills should be practised with older pupils, more willing to respond and discuss.

- Microteaching - or even a handout on skills plus videotaped models - would be valuable at beginning of course, particularly for students going straight out on teaching practice.

- Valuable/

- Valuable to start with microteaching, then assess use of skills in subsequent teaching practice.

- Microteaching particularly valuable for remedial work.

- Models should be based on normal-sized class.

- One microteaching session could be used for 'official' assessment of students.

- Microteaching must be used in a positive, encouraging way.

- Visual rather than audio feedback could lead to acute self-consciousness.

- Microteaching would not adapt well to some skills, e.g. disciplinary skills.

- "I would think an improvement of the course would be to have the microteaching lessons prior to going out on teaching practice, as it would be of immense help. As things stand just now, you are sent out on teaching practice with really no idea of what type of questions to ask."

- One of the most valuable teaching aids offered by the College.

- Microteaching in initial training could provide a more demanding and systematic approach than at present exists.

- Microteaching covers the middle ground between theory and practice.

(iii) Lecturers' questionnaire at end of session

Table XXXVIII

End of session questionnaire responses from History lecturers

Ref.	Question	Response	No Response
A1	Reactions to microteaching (n.b. a 'response' indicates broad agreement with one or more of the following comments)		
	- Effective teaching cannot be defined in terms of specific skills	-	
	- Specific teaching skills can be identified, analysed and consciously applied	7	
	- Theoretical analysis of skills is possible, but deliberate practice under controlled conditions is unlikely to improve classroom performance	-	
	- Specific skills can be strengthened by practice under controlled conditions	10	
	- A teacher must rely on personality and intuition rather than self-conscious techniques in developing his skills	2	
B1	The use of Appraisal Guide and Analysis of Skills helped to make teaching practice assessment		
	- more systematic		
	- agree	10	
	- disagree	2	
	- more precise and specific		
	- agree	11	
	- disagree	1	
	- more objective		
	- agree	9	
	- disagree	2	1
	- more reliable (i.e. same criteria applied to all students)		
	- agree	9	
	- disagree	3	
B2	Using the Appraisal Guide		
	- hindered task of balanced assessment		
	- agree	1 (+ query)	
	- disagree	9	1
	- made it difficult to attend to details of behaviour not specifically mentioned		
	- agree	1 (+ query)	
	- disagree	8	2

Table XXXVIII (Contd.)

Ref.	Question	Response	No Response
B3	Combining specific assessments (on question technique) and more general assessments proved		
	- fairly simple	1	
	- impossible to carry out adequately	2	
	- difficult, but not impossible	9	
B5	Factors which caused difficulty in achieving inter-tutor agreement over assessment of specific skills during group practice sessions (1971-72)		
	- difficulty in agreeing on concept of skill	3	
	- difficulty in noting and assessing a variety of skills	4	
	- difficulty in agreeing on concept of 'average' performance	8	
	- difficulty in agreeing on distribution pattern over 7-point rating scale	6	
	- difficulty in making reliable allowance for 'appropriateness' and 'grasping of opportunities'	8	
6 a	Reaction to criticism that Appraisal Guide "leads the observer into a no-man's land between scientific analysis and intuitive assessment"		
	- agree	3	
	- disagree	8	1 uncertain
b	Appraisal Guide obliges observer to take a teacher-centred view of lesson. More attention to pupil reaction and achievement would provide far more adequate assessment of teacher effectiveness		
	- agree	5	
	- disagree	6	1
7	Would you wish to continue regular use of structured form of assessment based on analysis of skills?		
	- Yes	8	
	- No	2	2 uncertain
C1	Should microteaching continue to be used in training of history graduates?		
	-Yes	9	
	-No	2	1 uncertain

Table XXVIII (Contd.)

Ref.	Question	Response	No Response
C2	(Assuming 'yes' to question C1) Should microteaching be incorporated		
a	- for all students at initial stage		
	- to train specific skills	5	
	- as a basis for general comment	3	
b	- for selected students at a later stage		
	- to give supportive training to those showing weakness in specific skills	8	
	- to give supportive training to students classified as 'generally weak'	5	
C3	(Assuming 'yes' to question C1)		
a	- How much time should any one student spend on microteaching?		
	- 2 days	4	
	- not sure	1	
	- as much as possible	2	5
b	- Should microteaching time be taken from		
	- teaching practice	7	
	- in-college courses	2	3
C4	Preference for microteaching sessions to be held		
	- in schools	4	
	- in college (involving transport of pupils)	5	3
5	Should microteaching (if introduced) be		
	- a shared responsibility among all lecturers	2	
	- responsibility of a small team	8	2
6	Assuming 3-phase graduate training, should microteaching be incorporated in		
	- Phase I	5	
	- Phase II	5	
	- Phase III	7	

General comments from lecturers' questionnaire:

A1 - Reactions to conceptual basis of microteaching

- Personality and intuition are not necessarily opposed to "self-conscious techniques". (5 comments to this effect)
- First comment would be more accurate if amended to read "...many different teaching styles, not all of which can be defined in terms of specific skills".
- "There are certainly undefinable qualities in many good teachers".
- Unorthodox means can be effective in achieving teaching objectives, but there is value in deliberate skills practice for
  - those lacking in imagination or inspiration
  - those whose enthusiasm leads to self-indulgence

A2 - Which skills should be added to or omitted from the Analysis of Skills?

ADD

- Quality of speaking (Clarity and Coherence)
- Reference to pupil reaction (Relationships)
- Teacher's ability to 'use his eyes' (Relationships)
- Identifying and anticipating sources of interruption and disaffection (Relationships)
- Delivery, i.e. voice, gesture, etc. (Clarity and Coherence)
- Avoidance of questions which oblige pupils to resort to guessing (Question Technique)
- Voice and personality (Clarity and Coherence, Relationships)

OMIT

- Choice of content (since student often has no choice)

B4 - Comment on suggestion that specific skills should be assessed for a short period of lesson, with remainder of lesson period devoted to general assessment.

- No means of ensuring that specific skill was practised during selected period, unless student was instructed to practise it; which would be an arbitrary procedure, making the lesson artificial and unbalanced. (8 comments along these lines)
- Any attempt at quantitative analysis would be most difficult during a normal classroom lesson. Tutor must concentrate on noting points of advice for the student, as well as making an objective/

objective analysis.

- A relevant analysis would be so complex that it would be almost impossible to apply. The lecturer is forced back on subjective assessment.

B5 - Factors contributing to difficulty in achieving inter-rater agreement in Appraisal Guide practice sessions.

- Difficulty in achieving a common standard of objectivity.

- Agreement on skills concepts may have been merely overt, concealing personal reservations.

- Discrepancies in lecturers' background of experience.

- Previous contacts with students seen on videotape may have influenced judgments in some cases.

B6a - Comment on the Appraisal Guide "leading the observer into a no-man's land between scientific analysis and intuitive assessment".

- "It depends a great deal on the observer - if he is willing to relinquish his pride in his own intuition he may be successful in his scientific analysis. On the other hand, if he places great store on human judgment, then he could be in a dilemma. With experience and perseverance, one could see the 'no-man's land' beginning to recede".

- This criticism could apply only to the skill of Relationships; but judgment here is based on classroom experience rather than intuition.

- Appraisal Guide removes temptation to give overall intuitive assessment which may fail to take account of major weaknesses.

- Assessment is definitely not an empirical science. Impossible to see any rules which would operate overall.

- The teacher's personality is a relevant factor. It could produce an effect opposite to that anticipated from a scientific analysis of the lesson.

- Doubtful about an analytical approach.

- Distinction between scientific analysis and intuition is improper.

- 'No-man's land' has a positive ring. A blend of scientific analysis and intuition can produce a valid and useful assessment.

B6b - Comment on criticism that Appraisal Guide obliges observer to take a teacher-centred view, whereas more account of pupil reaction and achievement would provide for a more adequate assessment of teacher effectiveness.

- a/

- A valid criticism. Relationships, as distinct from performance, can make or mar a lesson.

- The Appraisal Guide makes no provision for taking account of disparities in class discipline, etc. which can have marked effect on performance.

- Agree. A student could apply all the skills and yet be ineffective.

- What do 'reaction' and 'achievement' mean, and how can they be measured?

- Difficult in any case to make accurate measurement of 'effectiveness'. So much depends on the pupil variable.

- The tutor must concentrate on the student's performance, and in that sense his assessment must be 'teacher-centred'.

- Attempts to evaluate through pupil reaction and achievement would be too difficult.

C4 - Comment on major practical problems of introducing microteaching on a regular basis.

- There would be disruptive problems in transporting pupils to college; but microteaching in schools might give pupils the idea that they were merely guinea pigs.

- Problems would be

  - student numbers } (3 comments on these lines)
  - staff commitments }
  - sufficient time to put remedial/preventive techniques into practice
  - possible disruption if microteaching was carried out in schools

- Difficulty of approximating to a "normal" class situation.

- No insoluble problems.

- Tying up accommodation and the time of a staff member (to act as tutor)

- Expense of bringing pupils to college.

- Possible lack of student motivation arising from over-use of microteaching or concentration on skills which are relatively unimportant.

- Inter-departmental co-operation involved in releasing the students for microteaching.

C5 - Comment/

C5 - Comment on responsibility for microteaching: shared by all staff or responsibility of small team?

- Sharing has logistic advantages; specialism might prove more efficient.
- Better chance of achieving common standards if all lecturers took part.
- A small team, with one member changing each year, might be best.
- Volunteers only should take part.
- A specialist team of convinced tutors would be most effective and economical.

C6 - Comment on timing of microteaching (given a three-phase system for training of secondary teachers - phase I, initial college training; phase II, school experience; phase III, final college training).

- Microteaching could apply to all phases.
- General application of skills in phase I, concentration on specific skills on phase III, with selection based on reports from phases I and II.
- In phase I, microteaching could be used to bridge theory and practice. In phase III, teachers could learn from each other's microteaching performance.
- Delay microteaching until phase III, in order to put it into perspective and avoid getting 'bogged down' in practice of skills.

General comments.

- Impression from school visits that students seemed to concentrate on questioning even when inappropriate. Was this due to emphasis on microteaching? (This comment was echoed by another lecturer who felt that students seemed reluctant to abandon questions and answers in term 3, when they should have been experimenting with other approaches).
- Tutor criticism should be based not merely on what students did, but on what he <u>should</u> have done. Appraisal Guide made no provision for this.
- "Equally (referring to Appraisal Guide) I often found I could go through the seven points on questioning when what I wanted to say was, 'The student asked the wrong questions'".
- Microteaching can improve student confidence. They <u>feel</u> they are improving.

7. Interpretation, comment and conclusions

(a) Interpretation and Comment

Interpretation of the experimental results set out in chapter 6 must be preceded by a note of caution. One can legitimately point to indications, but it would be unwise to generalize too readily from them. Four limitations in particular must be kept in mind:

- a relatively small number of students was involved
- coding of behaviour was limited to one initial and one final lesson per student
- some of the skills which formed part of the microteaching programme were practised only rarely in the recorded lessons
- lesson content was not controlled by the experiment and in some cases was dictated by the school syllabus

Given these factors, it must be accepted that random behaviour and uncontrolled variables may both have influenced the results. For example, the researcher noted during the term 3 recordings that one or two members of the experimental group, with freedom to make decisions over content, selected topics which lent themselves to a great many higher order questions. Certain other students, who may well have had an equal grasp of question techniques, were obliged to teach topics which, by their nature, required more exposition and provided fewer opportunities for questioning at a higher cognitive level.

To set against these limitations, it should be noted that the design allowed for repetition of the experiment. Where the same pattern of significance occurs in each half of the term, it is reasonable to place greater confidence in the result and to be bolder in interpretation.

While the tests for significant differences between groups and terms were based on comparisons of percentage, the overall totals of questions/

questions set out in Table XIIIa-h provide some interest. The total number of questions increased from 1398 in term 1 to 1453 in term 3, but the difference is not significant; and the experimental group increase was slightly lower than that of the control group. It would have been disconcerting if microteaching had been associated with a marked rise in the number of term 3 experimental group questions, as happened in the study reported by Kelly (1973), where the experimental groups' rating rose to approximately five questions a minute, even though there was a preponderance of questions at the higher cognitive level. It is arguable that, beyond a certain point, an increased rate of higher order questioning, even in a discussion situation, is a sign, not of increased effectiveness, but of a failure to secure the kind of sustained and thoughtful responses which such questions should provoke.

Between terms 1 and 3, the number of Prompts increased only marginally overall, but both experimental and control groups increased their use of Probing, proportionately to the total number of questions asked. As far as is known, the control group students received no formal training in this technique; and it is therefore interesting to conjecture whether probing may be a behaviour developed almost unconsciously by student teachers as their confidence increases. It may be noted, too, that both experimental groups showed reductions in the proportion of questions attracting no response, whereas both control groups showed increases in proportions of this type of question. The indication - that increasing the proportion of higher order questions does not necessarily lead to an inability on the pupils' part to respond - is reassuring. The proportion of questions allowing no opportunity to respond dropped slightly for the experimental groups, but rose for the control groups, suggesting that, in the absence of corrective training, a gain in confidence may be associated/

associated with hurried questions which are either immediately repeated by the teacher or simply "thrown away".

Finally, it is noticeable that certain individual term 1 scores - for example, the very low $\frac{HOQ}{Q}$ scores by experimental student 1 in both the first and second half term groups - appear to be in marked contrast to the ratings of these students given by their tutors. It must be assumed either that the students' behaviour was inconsistent or that the tutors were influenced in their assessment by some factor - perhaps a trick of personality or style - which was not accounted for in the codings.

Table XIV shows the progress made, between terms 1 and 3, by experimental and control groups, on each of the categories under review. The first half term experimental group achieved a highly significant shift towards a greater proportion of higher order questions, with a corresponding reduction in the proportion of lower order questions. The use of probing questions also increased significantly. It was encouraging that the shift in favour of higher order questioning was accompanied by a significant decrease in the proportion of questions obtaining no response; but disappointing that the proportion of questions demanding no more than a yes/no response increased, even though the avoidance of this type of question had been a specific element in the microteaching training. The use of prompting increased and the proportion of questions allowing no opportunity to respond was reduced, although these changes were not significant.

The second half term experimental group achieved only one significant change, in reducing the proportion of lower order questions. The figures for higher order questioning were affected by what was perhaps a performance by the weakest student in the group, who in term 1 scored 54.5%. The performance on probing was also disappointing - the four students rated highest increased their proportion of probes, but the/

the remainder of the group did less well in term 3 than in term 1. It must be noted, however, that the term 1 performance of this group in probing was exceptionally high. Once again, the proportion of yes/no questions increased, though only marginally so, and the use of prompting decreased.

These figures must be viewed in contrast to the performances of the two control group sections. Apart from the improvement by the second half term group in probing, there was no significant shift in behaviour; and, indeed, performance of the first half term group deteriorated in four of the categories, while the second half term group performance deteriorated in six of the categories.

It was reasonable to suppose that teaching practice conditions in the second half of the summer term 1973 might have adversely affected the performance of both experimental and control groups. The latter half of any summer term is notoriously a difficult period for gaining steady experience in basic classroom skills. Moreover, the summer term of 1973 was unusually short, so that students assigned to schools in the latter half had only three weeks' practice, into which they had to fit lessons taught for tutors, as well as their recorded lesson for the microteaching experiment. However, a comparison of groups based on mean performances for each category does not suggest that the half term placing had any significant effect. In term 3, the first half term experimental group was superior to the second half term experimental group in seven out of eight categories; but in term 1 they had been superior in five out of eight. The position was reversed for the control groups, since the second half term group was superior to the first half term group in five categories (superiority having been equally divided in term 1). A comparison of first half term/ second half term performances based on the means of experimental and control groups combined shows that, in both terms 1 and 3, superiority

was/

was again equally divided.

Table XV provides evidence on the main hypothesis which the experiment set out to test: that "students given training through microteaching in sepcific skills will, in their subsequent classroom teaching, achieve performance in those skills significantly superior to the performance of a matched group who have been given no microteaching training." The lack of any significant difference between experimental and control groups in term 1 is consistent with the high inference ratings by tutors, which were used as a basis for matching. In term 3, the first half term experimental group was superior to the control group in all categories, with marked significance in terms of higher and lower order questioning and in probing. In the second half of the term, significant experimental group superiority was restricted to lower order questioning, though a measure of experimental superiority is shown in six of the other seven categories. Thus, in terms of significance for separate catgeories, the results fully support the main hypothesis in respect of lower order questioning; partially support it in respect of higher order questioning and probing; and do not support it in respect of prompting, avoidance of yes/no questions and reduction of questions allowing no opportunity for response.

The results of comparing experimental-control differences in gain scores, set out in Table XVI, confirm and extend the evidence of Tables XIV and XV. Between terms 1 and 3, the first half term experimental group made significantly more progress than the control group in shifting the balance of their questioning from the lower to the higher order level, and in reducing the proportion of questions which attracted no response. The second half term group made significantly more progress in reducing the proportion of lower order questions/

and of questions which gave pupils no opportunity to respond.

In weighing up the evidence of Tables XIV, XV and XVI, it is worth noting that, while the categories $\frac{X}{Q}$ and $\frac{HOX}{HOQ}$ cannot properly be said to represent skills, since they involved a variable (pupil ability) outwith the control of the students, they give an important indication of capacity to use other questioning skills appropriately and thus provide a link with the tutors' approach to rating. One would have expected that any group (for example, the first half term experimental group) which combined a significant switch in the direction of higher order questioning with a significant reduction in the proportion of questions obtaining no response would have been rated highly by tutors.

The graphical comparisons of group means of standardized scores, set out in Table XVII, provide a means of relating performances not only among groups but also among categories. It will be seen that, in general, experimental and control scores are relatively close in term 1 (as one would have expected, since the groups were matched) and tend to diverge in term 3. The exceptions are the scores for questions not involving a response, a category which was not taken into account in the matching procedure. Where specific categories are concerned, the divergence between experimental and control groups, from term 1 to term 3, is most marked for higher and lower order questioning. It will be noted, too, that term 1 scores for all four groups are below the standardized mean for higher order questioning and above the mean for lower order questioning. By term 3, the experimental group has reversed this situation, whereas the control group shows no significant change.

A study of individual standardized scores reveals no reason why the mean score of the second half term control group in prompting should have deteriorated so sharply, in contrast to the scores of the other three groups. The deterioration was shared evenly among nearly all members/

members of the group. A similar striking deterioration in the group's yes/no questioning performance is, however, explained by the fact that in term 3 one member achieved a standardized score of 6. The graph shows clearly the improvement, common to all groups, in the proportion of probing questions, the smallest gain being achieved by the group which did best in term 1. Both experimental groups reduced their proportions of questions not attracting a response and not allowing for a response, whereas the performance of both control groups deteriorated in these two categories.

The analysis of variance, detailed in Table XXa-h, was carried out to examine the interactive aspects of all possible sources of variance affecting student scores. In the event, the analysis revealed no fresh positive information. The treatment effect was shown to be highly significant in relation to the two main categories of higher and lower order questioning, and came near to significance in relation to probing and avoidance of questions allowing no opportunity for response. The half term effect was significant in relation to prompting, but came nowhere near significance for any other category. Examination of mean group scores reveals that the half term divergence in respect of prompting was due to the poor performance of the second half term control group, which returned the lowest score in term 1 and deteriorated in term 3, whereas the other three groups all improved in term 3. In terms of questions attracting no response (i.e. the occasion for prompting behaviour) group figures for the first half term (98 X questions) were almost identical to those of the second half term (97 X questions).

In view of the History department's concern with microteaching as a form of remedial training, it would have been helpful if the analysis of variance could have been used to test the interaction effect between treatment and initial ability on the selected skills.

Unfortunately/

Unfortunately, the design of the experiment, allowing for only one entry per cell, made it impossible to test this interaction. It is possible, however, to examine trends of experimental group performance as revealed by the data in Tables XIIIa, c, e and g. The question at issue is whether microteaching training had differing effects on students initially ranked lowest and highest on the skills concerned. Table XXXIX shows the experimental group term 1 to term 3 performance gains on those criteria where improvement was most significant (see Table XIV) namely

$\frac{HOQ}{Q}$ - first half term section

$\frac{LOQ}{Q}$ - first and second half term sections

$\frac{Pr}{Q}$ - first half term section

Mean performance gains are shown for

- the section as a whole
- students 1, 2 and 3 (ranking based on tutor assessment of term 1 lessons)
- students 7, 8 and 9 (tutor rankings)
- students 1, 2 and 3 (rankings based on coded term 1 performance on criterion in question)
- students 7, 8 and 9 (coder rankings)

The coder rankings (on the criterion in question) were used as a check on tutor rankings (based on sums of scores on the selected skills), in view of the low correlation known to exist (see Table XXIX) between tutors' and coder's term 1 assessments.

Table XXXIX

Trends in performance gains by experimental group students, term 1 to term 3

(a) $\frac{HOQ}{Q}$ - first half term section

	Term 1	Term 3	%age gain
Section mean	13.66	34.18	20.52
Students 1, 2, 3 (tutor ranking)	13.7	22.97	9.27
Students 7, 8, 9 (tutor ranking)	11.97	38.2	26.23
Students 1, 2, 3 (coder ranking)	28.1	41.33	12.23
Students 7, 8, 9 (coder ranking)	1.97	25.2	23.23

(b) $\frac{LOQ}{Q}$ - first and second half term sections

		Term 1	Term 3	%age gain
First half	Section mean	36.71	15.6	21.11
	Students 1, 2, 3 (tutor ranking)	29.73	20.2	9.53
	Students 7, 8, 9 (tutor ranking)	47.7	14.3	33.4
	Students 1, 2, 3 (coder ranking)	17.77	9.3	8.47
	Students 7, 8, 9 (coder ranking)	53.3	25.2	28.1
Second half	Section mean	41.97	17.86	24.11
	Students 1, 2, 3 (tutor ranking)	43.77	19.5	24.27
	Students 7, 8, 9 (tutor ranking)	38.7	14.83	23.87
	Students 1, 2, 3 (coder ranking)	23.3	13.4	9.9
	Students 7, 8, 9 (coder ranking)	58.13	18.77	39.36

(c) $\frac{Pr}{Q}$ - first half term section

	Term 1	Term 3	%age gain
Section mean	5.81	12.44	6.63
Students 1, 2, 3 (tutor ranking)	9.57	14.77	5.2
Students 7, 8, 9 (tutor ranking)	4.3	12.67	8.37
Students 1, 2, 3 (coder ranking)	13.1	15.83	2.73
Students 7, 8, 9 (coder ranking)	1.37	10.67	9.3

In interperating these figures, it must be borne in mind that, because of the apparent unreliability of tutor ratings and lesson-to-lesson variability among students, correlation for any group between term 1 and term 3 performance, whether measured in terms of tutor or coder rankings, would be far from perfect. In consequence, the term 3 mean performance for students ranked 1, 2, 3 and 7, 8, 9 in term 1 would show a regression towards the overall group mean. Since, for each of the criteria reviewed, group mean performance rose markedly between terms 1 and 3, the result would be a small mean gain for the students initially ranked most able and a much larger mean gain for the students initially ranked least able. However, if term 1 to term 3 gains had been skewed in favour of the initially more able students, the regression effect would have been reduced. The inference to be drawn from Table XXXIX is therefore that there is no contra-indication from the performance data to the use of microteaching as a remedial technique. If one examines the coder ranking figures, it will be seen that, on each criterion, the students weakest/

weakest in term 1 were still weaker in term 3 than the students who were initially strongest; and, with the exception of the $\frac{LOQ}{Q}$ criterion in the second half of the term, the performance achieved by the weakest students in term 3 was still inferior to the term 1 performance of the best students. However, the performance gap between the most and least able students had in all cases been considerably reduced. These apparent trends are of interest, since on the one hand, large institutions such as Jordanhill College tend to look on remedial training as the most "practical" way in which microteaching can be used; and, on the other hand, the evidence on self-confrontation through video playback collected by Fuller and Manning (1973) leads them to conclude that "there is the danger that 'poor' teachers have characteristics which make them least likely to benefit from the treatment, but more likely to be treated in the hope that they will see what they are doing wrong and correct it." It may be relevant that self-confrontation in the Jordanhill experiment took the form of audio rather than video playback and that the subjects were relatively mature students. It was the impression of the researcher that the students who initially performed indifferently on the skills were characterized not so much by anxiety and low self-esteem (the kind of qualities which, according to Fuller and Manning, produce bad reactions to self-conforntation) as by a sturdily traditional view of the teacher's role or an excess of zeal about imparting information.

The experimental v. control comparisons of superiority shown in Tables XXI - XXIII provide a supplementary way of examining the significance of differences between groups. Even where specific category differences may not be sufficiently large to achieve significance as measured by t-tests (see Tables XV and XVI), the overall trend of superiority may itself be significant in terms of probability. Table XXIV shows that, on this basis, there was no significant difference between/

between experimental and control groups in term 1, but that, by term 3, significant differences had emerged for each half term. This significance characterises comparisons both of term 3 performance and of term 1 - term 3 gains.

Tables XXV and XXVI set out the percentages of teacher talk recorded throughout the experiment. It will be seen at a glance

- that the range is wide, from 46% to 93%
- that the mean percentages are high (the overall mean is 73.65%)
- that there is no significant correlation, either positive or negative, between percentage of talk and tutor rating of student performance

The mean percentages are in fact in excess of those reported over the years by other researchers, for example Stevens (1912), approximately two-thirds; Corey (1940), 64%; Bellack (1966), 72.1% of lines of transcript. Table XXVII shows that no group improved significantly between terms 1 and 3. Table XXVIII indicates that, while both experimental groups showed initial superiority over their matched control groups (the superiority being significant in the first half of the term), the experimental - control differences increased slightly in favour of the experimental groups in term 3.

These results are in marked contrast to those obtained by Borg (1970) in his testing of Minicourse 1: Effective Questioning- Elementary Level. The teachers involved in his main field test scored a pre-test teacher talk percentage of 51.64, and this was reduced in the post-test to 27.75. It would appear that the "teachers talk, children listen" tradition was too firmly established in the students' minds to be broken by a brief research project; though it must be remembered that no teacher can force sustained responses from pupils who have developed a deeply engrained habit of verbal parsimony. For this reason, it was probably wise not to pursue the original intention, which appeared logical in theory, to use the/

the percentage of teacher talk as the principal measure in assessing the effectiveness of the microteaching training programme. The measure has the attraction of high reliability, but its validity is questionable. The percentage of teacher talk is inevitably affected by a number of variables which would be very difficult to control, for example

- the ability of the pupils. A student's progress would be hidden if his term 3 class were less able - and less fluent - than his term 1 class.

- the interest of the subject-matter. This will very probably influence the pupils' readiness to talk, but it would be difficult for a student on teaching practice to predict the subjects which would interest a particular class.

- the degree of pupil knowledge; again, difficult for a student to predict in relation to a specific topic.

- the social atmosphere of the school. However able, a student teacher is unlikely to obtain much response from a class with a long tradition of inglorious muteness.

Moreover, the measure has certain inherent defects as an indication of effectiveness. For example,

- it is arguable that a teacher who talks a lot but who talks well may be more effective than a tongue-tied teacher who pauses as much from inadequacy as from a positive desire to stimulate response in the pupils.

- a teacher who tries hard to distribute his questions widely/

widely may obtain a higher proportion of brief responses than a teacher who concentrates the discussion on a few bright pupils, ignoring the rest of the class.

- a teacher who asks vaguely worded questions, or questions which are too difficult for the class, may spend a long time pausing for response. This kind of pausing is not necessarily an indication of effective teaching.

Measurements of teacher talk were certainly worth taking into account in assessing the impact of microteaching; but, for the reasons stated above, they would have been suspect as a basis either for initial matching of students or for making final judgments on training effects.

One of the subsidiary aims of the research study was to examine the extent of agreement between assessments of students made by tutors on school visits and assessments of the same students based on a quantitative analysis of their teaching behaviour. Since experimental and control groups in each half term were composed of matched pairs, their rank order, based on tutor ratings, exhibited perfect positive correlation. Assuming inter-tutor reliability, congruity between the elements of behaviour assessed by the tutors and those measured by the coding instrucment, and consistency of performance by students from one lesson to another, the correlation between tutor rank orders and coder rank orders should have been high. In fact, Table XXIX shows the correlations to be very low, and it must be assumed that one or more of the following factors were operative:

- the reasonably high level of inter-tutor reliability in assessing specific skills performance, achieved by the end/

end of the preparatory year, had been lost in the following session.

- in making their assessments, tutors were taking into account elements of behaviour and style which were not covered by the quantitative measures of performance and which had a varying effect on tutors' judgments when compared with the coder's measurements.

- students were displaying wide inconsistencies of teaching behaviour from one lesson to another.

Rank order correlations based on recorded lesson performance by experimental and control groups in term 1 were carried out to discover whether the groups could still be said to be matched in terms of the coder's assessments. Table XXX indicates that, while there was some measure of correlation, this was not at a significant level. The evidence of Table XXXI is important as an indication that rank order correlations based on successive assessments by tutors in term 1 were also too low to achieve significance. Since each student was assessed on each occasion by the same tutor, the question of inter-tutor reliability does not arise. The relatively low correlations must be due either to student inconsistency in teaching behaviour or to tutor inconsistency in assessment. It will be remembered that the tutor ratings used as a basis for matching were the averages achieved from three assessments on separate term 1 visits. It would appear that these averages concealed wide variations in assessment from one visit to the next. The reliability of the tutor ratings is neither proved nor disproved by these variations, or by the low correlations with the rankings based on coded measurements, but it must be called in question.

Another/

Another subsidiary aim of the study was to gather evidence of the effect of microteaching training on students' classroom performance as a whole, and it was to this end that the tutors' Appraisal Guide was designed to cover general teaching ability as well as specific questioning skills. Table XXXII shows that no significant differences between experimental and control groups emerged, either in term 1 or in term 3, from the data provided by the Appraisal Guide on overall teaching competence. For each half term, the initial differences between groups increased slightly - in favour of the experimental group in the first half term and the control group in the second half term - but not significantly so. There is thus no indication that, in the tutors' judgment (which was made in ignorance of which students were in the experimental group and which in the control), microteaching had any significant effect, either beneficial or adverse, on teaching performance as a whole. This finding is of interest, in view of the impression formed by some tutors (reported below) that microteaching was distorting the balance of teaching behaviour as the year progressed.

Table XXXIII relates to the ratings by tutors of performance on specific questioning skills in term 3. Again, it must be stressed that these ratings were made in ignorance of which students had received microteaching training in the skills concerned. Only one difference achieves significance, but the recurrence of control group superiority - for six out of seven skills in the first half of the term and for all seven skills in the second half - is itself significant. These findings must be contrasted with the evidence of Table XXII, which summarizes the coder's findings of superiority - all eight categories favouring the experimental group in the first half of the term and seven out of eight categories showing experimental group superiority in the second half term. The only category where there is a measure of agreement between tutors and coder is the use of higher order questions for the first half term group; and even here the small difference recorded by the tutors

(t = 0.532) contrasts with the highly significant differences recorded by the coder (t = 6.588).

These discrepant judgments made it necessary to proceed with tutor rating of the discussion section of the term 3 recorded lessons, to discover whether differences between coder and tutors were due primarily to inconsistencies in student performance between one lesson and another. Table XXXIV shows clearly that only one pair of tutors (they were, in fact, the principal and senior lecturers in the department) reached any consistently significant level of agreement among themselves. Agreements within the other four pairs of tutors were low overall and fluctuated widely from one category to another. The only conclusion that can be drawn is that, with the notable exception of the two senior members, tutors' ratings of student performance on specific teaching skills had become highly unreliable after a lapse of eighteen months since the final reliability trials of the Appraisal Guide. This finding supports the evidence of earlier research studies (for example, Biddle, 1962; Cicirelli, 1969; Oppenheim, 1970) concerning the unreliability of assessments based on subjective ratings. The evidence of unreliability among the Jordanhill tutors makes it unwise to pursue the implications of apparent differences between high and low inference assessments. Table XXXIV reveals no consistent pattern relating levels of inter-tutor agreement to levels of agreement between each tutor and the coder. The figures in column 5 carry some suggestion that, where the tutors are in consistent agreement, there is some measure of agreement between them and the coder; but, even here, the pattern breaks down for prompting and probing, which is puzzling in view of the fact that these skills had been precisely defined in the recorded lesson rating form.

The lack of inter-tutor reliability means that little if any significance can be attached to the comparisons of superiority based on the mean ratings of/

of tutor pairs and set out in Table XXXV. Nevertheless, it is of some interest that the pattern is in marked contrast to that shown in Table XXXIII. In so far as any reliance can be placed on the tutors' findings, the shift in superiority would suggest that the experimental group students had retained from their microteaching experience a capacity to use the skills when the occasion (in this case the term 3 recorded lesson) demanded, but that they were not necessarily employing them to the same extent in their normal teaching routine.

Data from the questionnaire completed by the experimental group of students in term 2 (immediately after microteaching) and term 3 (after the final teaching practice) are summarized in Tables XXXVI and XXXVII. In general, reactions are favourable, but the high incidence of thoughtful and perceptive comments suggests that students went beyond a mere careless acquiescence in the experiment. The term 2 questionnaire had three main sections: on the handout ("Four Questioning Skills"), on the videotape models and on the actual experience of microteaching. Response to the first section indicated that a large majority of students had taken the trouble to read the handout at least twice, and had found it clearly set out, although opinion was divided as to whether its advice was "too theoretical". The fact that no student, by the end of term 2, was prepared to disagree with any of the handout statements suggests that at least the description of questioning behaviour in terms of component skills appeared relevant, in retrospect, to the students' own experience. The modelling of the skills on videotape appeared "reasonably useful" in making the nature of the skills clearer to a majority (15 students) of the group. The students' timetable made it necessary to compress each cycle of microteaching into one day, with the viewing of the model tape immediately preceding the start of the practice, but it is interesting that eleven of the group would have preferred a gap of a day or two between viewing and practice, so that they could assimilate the advice

of/

of the model into their own lesson planning. The comments (B6) on the videotape models are also of interest. The cueing and reinforcement provided by the commentary links would seem, to a majority, to have strengthened the message of the teaching sequences. Opinion appears to have divided between those who wished the models to have a micro-teaching format and those who looked for an illustration of the skills being applied in a normal classroom setting. (In fact, the models were a compromise, since they were recorded in a studio classroom with a group of about ten pupils). Clearly, a "microteaching model" and a "classroom model" have different functions, and perhaps both have a place in a full cycle of instruction.

Response to section C (Microteaching Practice) indicated that the actual experience of microteaching was found on the whole to be acceptable, though a number of students noted the "artificiality" of teaching a small group of pupils for a brief period. It is reassuring that only one student was distracted by the recording equipment, and that no-one confessed himself confused about the nature of the skills which he was practising. Peer grouping at the review stage was found positively useful by fourteen students (this supports the findings of Turney, Owens and Hatton (1970) with Australian students) and no-one expressed a preference for reviewing performance with the tutor alone. The conclusion by McIntyre (1971) that "the provision of tutors appears necessary" would seem to be fully borne out by the comments of the Jordanhill group. Fourteen of them wished to have the tutor present at least at the review stage and an equal number found his comments "very helpful". There are hints in the comments on the tutor's role (3.C.3) that support the suggestions of Johnson and Knaupp(1970): that the students wish the tutor to provide technical assistance ("putting his finger more clearly on particular points"), but to allow opportunity to find their own teaching style (e.g. fear lest a history department tutor might prove "inhibiting"/

"inhibiting" to sensitive students). The use of evaluation forms for completion by students at the review stage met with a mixed reception - apparently they were a help to some, a distraction to others. Perhaps the sensible course in future experiments would be to make the forms available to those who found them a help. Finally, the reaction to the re-teach lesson, as a means of improving confidence and correcting initial weaknesses, was good. The subjective impression of tutor and researcher, however, concurred with the suggestions of Orme (1966) and Skailand (1972) that the re-teach stage appeared to add little to the students' competence in applying the skills being practised.

The final section of the questionnaire, seeking general comments, produced the most considered responses. The suggestion that some guidance should be given in selecting material appropriate to the practice of the selected skills, and that the organizers of microteaching sessions should be alert to the dangers of tedium and overwork are, in the opinion of the researcher, valid. The comments on developing confidence to pause for probing and re-directing, and on the likely reaction of departmental heads if these techniques are employed, reflect two sides of a single coin. They recall the experience of Brusling (1974), on the possible clash between training values and teaching goals, and suggest that microteaching at the pre-service stage may be fully effective only if it is complemented by an organized programme of in-service debate on the implications and validity of technical skills training.

The term 3 questionnaire gave students the chance to make a retrospective assessment of microteaching, in the light of subsequent teaching practice. It was encouraging that all the group claimed to have made at least occasional attempts to apply the skills during their term 3 period in schools. Their perception of the difficulty or success they experienced should be compared with the analysis of their actual use of the skills in term 3, set out in Tables XIII e and g, and the measure

of/

of their term 1 to term 3 performance gain, shown in Tables XIV and XVII. Restricting "one-word-answer" questions and making optimum use of higher order questions - skills in which the majority of the students reckoned to have achieved moderate or good success - relate to the $\frac{LOQ}{Q}$ and $\frac{HOQ}{Q}$ ratios where the coding showed significant gains. But the students were also confident about their success in restricting questions demanding no more than a yes/no response, and here the coded lessons showed gains of no significance. Pausing for response was felt by the majority of students to be a difficult skill to master, although the coding indicates some gain in the related $\frac{n}{Q \text{ incl. } n}$ ratio. The only other discrepancy between perception of performance and coded behaviour was in prompting, where the second half term group considered itself "moderately successful", although the coded lessons showed a slight deterioration in performance from term 1 to term 3. The reasons given for difficulty in applying the skills indicate an understandable preoccupation with disciplinary problems related to class size and ability level, and with the practical problem of "getting through" the material which they were expected to cover. The impression, noted by several of the students, that pupils were unused to being questioned at a higher cognitive level and were undisposed to take such questioning seriously suggests once again the need for close co-operation between those responsible for college training programmes and the senior members of school staff involved in supervising student teaching practice. On the one hand, there should be a concern by the college that the selection of skills is relevant to the teaching situations in which students will be involved. On the other, there should be an understanding within the schools of the approaches advocated by the college; or at least a willingness to consider them, rather than reject them out of hand as theoretical and unrealistic. This belief held by all members of the group that the application of these particular questioning skills had a beneficial/

beneficial effect on the length and quality of pupil responses indicates that the training programme had, in fact, some relevance to their task as teachers. Finally the group's conviction that transfer of the skills from the microteaching to the full-scale class situation was possible, even though difficult, relates to the main hypothesis being tested by the experiment.

Section B of this questionnaire sought the students' views on the value of microteaching as an element in the graduate training course. The reaction is clearly favourable, with all members of the half term groups recommending its retention in future sessions. It would appear from the response to B3 and from the general comments that the students saw a use for microteaching both initially and at the midpoint of the course; and that while they considered it valuable as a form of remedial training, they were anxious that all students should experience it at some stage. There is a strong preference for a complete cycle of training; an interesting reaction, in that some college departments have tended to prefer a restricted form of microteaching, without going to the length of defining and analysing skills or of running a complete teach-reteach sequence. The statement that microteaching could provide a more demanding and systematic approach was echoed by several informal comments. It appeared that a number of graduates welcomed the element of analytic rigour which they associated with the microteaching programme.

Since the development of any future activity in microteaching depended on lecturers' acceptance of the techniques, it was essential to analyse the reactions of History department tutors in the detail set out in Table XXXVIII. The questionnaire was completed by twelve members of staff, including the senior and principal lecturers. The response to Section A (conceptual basis of microteaching) was clearly positive, though the belief in/

in the value of analysing and practising specific skills was matched by an insistemce on the importance of personality and intuitive approaches, which, as several tutors pointed out, did not necessarily exclude skills training. It is interesting that four of the seven suggestions for additions to the Analysis of Skills come under the heading of Relationships, implying once again a concern with personality as a means of exercising effective control.

Section B of the questionnaire, delaing with the use of the Appraisal Guide, sought the lecturers' reaction to a more systematic and detailed form of assessment than they had been using before the start of the experiment. On the face of it, the majority reaction is one of approval; a belief that the Appraisal Guide helped to make assessment more systematic, precise, objective and reliable, that it did not hinder the task of balanced judgments and that its dual function (combining general and specific assessments) made the lecturers' job difficult but not impossible. There was no support for the suggestion of an alternative tactic of assessment (taking a short section of a lesson and applying a specific analysis to it), which was made with the Stirling Lesson Sampling Instruments in mind. These positive reactions to the Appraisal Guide must be set against the evidence of the lecturers' failure to maintain reliability in using it. This failure was attributed by the lecturers (see B5) mainly to difficulty in agreeing upon a concept of "average" performance and in making reliable allowance for appropriate use of skills. It is hard to see how these difficulties can be overcome, other than by prolonged and intensive training based on analysis and group discussion of recorded lessons. Questions 6a and b were designed to probe more deeply into the problems of assessment raised by the use of the Appraisal Guide, and the lecturers' comments raise interesting issues. One can detect a division of opinion between those who welcome some measure of control over subjective reactions and whose who question the/

the validity of assessing effectiveness in terms of quantitative measures of the use of skills; and between those who would attempt an evaluation of pupil reaction and achievement and those who would concentrate on observing the teacher's behaviour. Behind this divergence, it is possible to sense differing views of the tutor's role vis-a-vis the student on teaching practice. The Appraisal Guide was designed to produce a relatively objective evaluation of behaviour. But the tutor is not simply an analyst; he is conscious of a pastoral role, which involves separating goats as well as feeding sheep. It could be argued that both these functions could legitimately derive from a prior analysis of behaviour; yet there are hints, here and elsewhere in some of the lecturers' comments, that they viewed the completion of the Appraisal Guide as a task additional to their "normal" job of evaluation and assessment, rather than a basis on which judgments could be built and from which counselling could be developed. It is perhaps significant that, although (in B7) a majority expressed a wish to continue the regular use of some structured form of assessment, there has been, as far as the researcher is aware, a reversionsince the end of the experiment to the traditional tutor's report and general competence rating on a five-point scale.

As for Section C (future developments of microteaching), the lecturers agreed (though not unanimously) with the students that microteaching should continue to find a place in the graduate training course. Question 2 was phrased partly to test the extent of interest in the component skills approach to microteaching. It would appear that the majority of tutors wished to base this type of training on specific skills and to use it for remedial purposes. In relation to question 3 (time to be given to microteaching), it was noticeable that no one made any comment suggesting that microteaching should form part of an analytic approach involving elements of both in-college course work and teaching practice.

The/

The comments on C5 again indicated a division between those who were enthusiastic about the potential of microteaching and those who were still dubious, though not actively hostile.

There were a number of more general comments scattered throughout the completed questionnaires. The impression of the two lecturers who felt that the experiment had resulted in an undue concentration on questioning is not supported by the coding (Table XIII shows no significant term 1 to term 3 changes in the overall totals of questions) or by the students' assertion (Table XXXVI, A3 and A4) that they referred back to the Questioning Skills handout only occasionally, if at all, during the session. In fact, only a minority of the history graduate population took part in the study, either as experimental or control subjects, and it is improbable that the remainder took the trouble to refer back to the handout after the initial week of term 1. As for the problem of the students who asked the "wrong" questions, it is true that the selected questioning skills (which were concerned with measurable behaviour) did not cover this. The problem is complex, since a question may be "wrong" in respect of content, phrasing, timing or content. It is doubtful whether a talent for asking the "right" question at the right moment can be classified as a skill at the level defined (in 1974) by McIntyre ("a deliberately acquired, habitual pattern of teaching behaviour"), though it is clearly an element in effective teaching. Finally, it is interesting that the comment on the power of microteaching to increase student confidence was made by the tutor who had acted as supervisor throughout the term 2 sessions. His belief finds support in the comments of those students who, after their final teaching practice, noted the confidence which microteaching had given them to control the pace and level of their questioning.

(b) Conclusions

Rvidence supporting the main hypothesis is summarized in Tables XV (Experimental - Control group term 3 comparison), XVII (group means of standardized scores), XX (analysis of variance), XXIV and XXXV (probability factors associated with experimental group superiority in term 3). In the light of this evidence, it can be claimed that, at least on the basis of the coder's assessments, the experimental groups, on a post-microteaching classroom teaching occasion, made significantly greater use than did the control group of some but not all of the questioning skills in which they had been trained; and that their overall use of these skills was, in terms of probability, significantly superior to that of the control groups. These findings hold good for both half term sections, though the superiority is more clearly marked in the first half of the term. Any claim that the results of the study establish the main hypothesis must be subject to a number of qualifications. The danger of generalizing from a small experiment restricted to a single subject area has been stressed at the beginning of this chapter. One cannot say whether the same results would be obtained with different skills related to other specific specialisms. Moreover, the evidence merely indicates that, on a particular occasion which was known to be related to microteaching training, stduents showed a capacity for transferring acquired skills to classroom practice. There is no objective evidence to show that this transfer was habitual, though the students' claim (see Table XXXVII, question A1) that they made conscious attempts to apply the skills at least occasionally must be taken into account. But having made these qualifications, it must be stressed that the results provided a reasonable measure of support for the belief that some element of mciroteaching may usefully be incorporated into the pattern of the graduate teacher training year. For one thing, the incidence of experimental group superiority was to some extent replicated/

replicated in each half of the term. It must also be borne in mind that the performance gains were achieved after only two days of microteaching, allowing for one complete cycle of training on each of the skills; half the minimal period considered by the Stanford University research workers to be essential for effective training. Under these circumstances, it was perhaps not surprising that the stduents did better in acquiring and retaining the major skills of shifting emphasis to higher cognitive questioning and of probing questions - skills whose employment involved an element of conscious decision - than they did in avoiding yes/no questions and questions giving no opportunity for response: patterns of behaviour which may well have been strongly linked with habit or personality. Finally, in view of the limited and inconclusive amount of research evidence on transfer of microteaching performance gains to classroom practice, it is arguable that any study such as that at Jordanhill which provides at least some positive indication of transfer is worth following up. In particular, the fact that significant differences between experimental and control groups in classroom performance were found approximately three months after the training period provides more grounds for optimism than the evidence of Brusling and Stukat (1972), who found that positive effects of microteaching training had largely disappeared two months later when the students were observed in regular classrooms; and of Copeland and Doyle (1973), whose experimental group, seven weeks after they had completed an extensive period of microteaching, showed no significant superiority over a control group, on the basis of coded classroom lessons.

The second aim of the research study was to examine the feasibility of identifying and analysing a set of skills appropriate to history teaching in secondary schools, and of creating an appraisal instrument, based on these skills, which could be used reliably by lecturers. Like any well stocked rag-bag, the "Analysis of Teaching Skills", produced in the

first/

first year of the project, had the merit of being comprehensive. The rationale of this analysis has been described in chapter 3, section a, where reference was made to the difficulty of agreeing upon the concept of a teaching skill and to the decision to distinguish three categories of basic, composite and specialist skills. Clearly, teaching, as a form of social interaction, involves a hierarchy of skilled behaviour, in the sense employed by Argyle (1969), where "the samller, lower-level units in the hierarchy are more habitual and automatic than the larger, higher-level units". It therefore seemed logical to organize the analysis under the headings of a number of broad competences, each of them sub-divided into "lower-level units". In retrospect, the problem implicit in the analysis would appear to be that, under each broad heading, and as between the different headings, the specified units of behaviour varied constantly in level, so that the reader could retain no consistent concept of the nature and scope of a "sub-skill". Thus, some skills - for example, Question Technique 9: "The teacher acknowledged good answers with approval and poor answers with tolerance" - are at the relatively low level implied by McIntyre's definition of a teaching skill as "a deliberately acquired, habitual pattern of teaching behaviour which, in specified types of context, tends to be effective in achieving objectives of a specified pattern". Other skills - for example, Varying the Pace and Rhythm 6: "The teacher balanced periods of intensive effort and activity against quieter, more relaxed periods" - are more broadly based and at a higher level, characterized by selective perception of cues, anticipation, decision making and response to feedback. Again, it is difficult to equate the planning skills, listed under Preparation, with the social interaction skills listed under Presentation, Pupil Involvement and Relationships. These conceptual uncertainties in the Analysis of Skills may account in some measure for the difficulty experienced by some lecturers in achieving and maintaining reliability in their use of the Appraisal Guide. It may well be that the task of constantly/

constantly changing perspective in their view of the lesson was more than they could cope with; and it is significant that the Appraisal Guide reliability trials revealed a much higher level of inter-rater agreement for assessments of broad teaching competences than for judgments of specific skills performance. However, in the tutors' own view (see Table XXXVIII, B5), the main problem in achieving reliability with the Appraisal Guide lay not so much in defining skills as in agreeing upon the concept of average performance and in making allowance for "appropriate use" and "grasping of opportunities". It must be admitted that the research project found no solution to the difficulty of devising a systematic technique of appraisal which could accommodate high inference assessment while achieving reliability at a departmental level. The task of identifying skills was a necessary first step in any process of reliable appraisal; but it must be concluded that, at the end of the day, the main value of developing such a comprehensive analysis and such a detailed appraisal instrument lay, first, in the opportunity for communal exploration of what had been to some extent private territory; and, second, in the exposure of hitherto unacknowledged problems inherent in subjective judgements.

The third aim of the inquiry was to discover the extent of agreement between assessments of students made by tutors on school visits and assessments of the same students based on quantitative analysis of their recorded teaching. Table XXIX makes it clear that no such agreement was achieved and Table XXXIVindicates that the lack of agreement was associated with an inability among tutors to make reliable assessments of specific skills. Because of this unreliability, it is not possible to infer that there were factors implicit in high inference judgments which pointed in a different direction from low inference measurements of performance. Moreover, an important question on which it had been hoped to gather evidence remains unsolved: the question of the link between increased/

increased use (fostered by microteaching training) of specific technical skills and teaching "effectiveness". The Jordanhill lecturers were convinced that the key to the effective use of skills lay partly in appropriateness - the right skill at the right moment - and partly in teaching style and personality. Their view is supported by the findings of Rosenshine and Furst (1971), that the presentation variables correlating most consistently with pupil achievement were identifiable through high inference ratings of performance. Conversely, some of the lecturers shared the doubts expressed by St John-Brook and Spelman (1973), that, through the component skills approach of microteaching, "highly visible but relatively trivial teaching skills have been emphasized at the expense of more subtle and individual techniques" and that "the emphasis on individual skills may hamper student teachers in developing a full response to the highly complex patterns of behaviour which can be seen in the real classroom". Evidence of term 3 classroom performance, based on demonstrably reliable high inference assessments, would have provided a first step towards establishing a relationship, positive or negative, between the kind of training offered by the microteaching programme and increased effectiveness. As it is, no conclusion can be reached and there are no clues to follow up.

Research during the 1960's at Stanford University and at the University of Illinois had shown that, within an experimental group, microteaching performance was a good predictor of subsequent performance within the classroom, but no major investigation could be found of the effects (beneficial or adverse) of microteaching on classroom teaching as a whole; though the study by Britton and Leith (1971) suggested that the effect of such training might be beneficial. The question was of particular concern in the context of graduate training, since it was feared that, within the limits of the graduate training year, a programme of intensive instruction on specific skills might lead to unbalance/

imbalance. Table XXXII indicates that, comparing experimental and control group means, there was no significant relationship, either negative or positive, between microteaching and subsequent overall teaching competence. To supplement this finding, a comparison was made between the tutors' ranking of the experimental group students, based on their overall teaching marks in term 3, and the coder's ranking of the same students, based on the sum of their rankings in specific questioning skills. Correlation coefficients were, for the first half term group, -.008 and, for the second half term group, zero. The importance of these comparisons lies in the fact that tutors' ratings of overall lesson performance had been shown (see chapter 3, Tables V and VI) to be consistently reliable. The conclusion to be drawn from this apparent lack of any connection between microteaching and general competence is reassuring, in that lecturers' fears of possible adverse effects find no support from their own assessments; but disappointing, in that a significant experimental group change (as measured by the coder) in an important aspect of teaching behaviour either did not extend into lessons taught in the presence of tutors or, if it was indeed present, went unnoticed by the assessors.

As the research study developed, it became apparent that no final conclusions could be reached about the practical problems involved in organizing microteaching within the graduate training year. In some ways, the element of research added to these problems, for example by introducing the need to maintain confidentiality and by imposing a strict timetable of pre- and post-testing. In other ways, the researcher carried out most of the organizational work (liaison with schools, transport, booking of accommodation, etc.) which would normally be the responsibility of the department concerned. From the students' point of view, the only practical drawback to microteaching lay in disrupting a sequence of teaching practice and (for some students) in extra travelling/

travelling between home and college. The lecturers, in completing their questionnaire (see Table XXXVIII, Comments on C4), saw the major problems as student numbers and staff time. For an experiment of this size, there was no real problem of obtaining pupils from nearby schools; but transporting them was a time-consuming and expensive business, and accommodation (two adjacent rooms are required for even a restricted microteaching session) was always at a premium in the college.

More direct evidence on the kind of practical problems likely to be faced in future microteaching developments was obtained by the History department in the year following the research experiment. It was decided to continue with microteaching on much the same lines. The tutor who has acted as supervisor in the research study took on the job of organizer and submitted a full report at the end of the session. This showed that offering two days of microteaching to all graduate students in the department would have involved a tutor from 9 a.m. to 5.15 p.m. for four days of each week during an entire term. In the event, two days' training was given to thirty students during the second term, taking up twelve days in all. Apart from the supervisor, two other lecturers were involved in ferrying pupils to and from schools and in looking after them during training breaks; and an audio-visual technician operated the audio recording equipment throughout each day. Training was concentrated within term 2, which was a logical decision from the History department's point of view, since it enabled them to base their choice of students on the evidence of term 1 teaching performance and to allow the students to consolidate their microteaching experience during their term 3 school practice; but, for all students studying a subsidiary subject in addition to History, approval for release to microteaching had to be obtained from the subsidiary subject department; and achieving co-operation presented problems in some instances.

Bringing/

Bringing students into college from teaching practice can involve organizational problems perhaps more typical of the college graduate training year than of the pattern of university work. The History department's experience was aptly described by the tutor:

> "The first group of five students, eagerly awaited by a shivering tutor on the first Tuesday in January, suffered the 'Little Nigger Boys' syndrome. One had suffered a heart attack before Christmas and was still absent. One transferred to Moray House with effect from 1st January. And so we had three. But not for long. One student heard of a domestic crisis at lunchtime and did not reappear then or the next week. And so we had two. Except that the next week Scotland took its turn to enjoy a rail strike, and one student telephoned from Airdrie to say that it was impossible to get on a bus. And so we had a 'group session' of one student."

In replying to a questionnaire on this experience of microteaching, a number of students stated that they would prefer it during in-college times, to avoid making their teaching practice disjointed. According to the tutor, the highly complicated timetable, providing different patterns not only for the various subject groupings but also for Honours and Ordinary graduates, made it impossible to gather a "readily definable group" together for microteaching during the in-college periods. Other practical problems mentioned by the tutor were

- the fact that, "for a group of five students to do microteaching, you affect the timetables of possibly three lecturers"

- the difficulty of co-ordinating microteaching with teaching practice. "To some extent, tutors are aware of the names of those who do microteaching, but it is purely a matter of chance whether the microteaching tutor sees the student again in school."

The/

The tutor's summing up in his report is not optimistic:

> "In the light of timetable and manpower difficulties outlined above, it is likely that we must concentrate on students in need of special help and take up only two days each half term. With groups containing only weak and poorly motivated students, the benefits may be minimal in terms of improvement in performance."

It can be seen that practical problems of this kind are characteristic of an attempt to fit microteaching as an "extra" into an established time-table and curricular pattern. Given this approach, the conclusion must be that a minimal experience of microteaching can be organized (though with some difficulty) for only a minority - perhaps a quarter - of students in a department such as History, with ten to a dozen tutors, handling approximately 120 graduates as well as large numbers of Diploma students and a B.Ed. group. It would seem impossible, in practical terms, to offer microteaching to all graduates without major re-planning of the curriculum as a whole.

The final aim of this study was to examine the reaction of students and lecturers to the experiment. These reactions have already been commented on in some detail. In conclusion, it would seem worth emphasizing two points which have a bearing on any future development of these training techniques:

a) the students' reaction

- the recurring suggestion in responses to the final questionnaire that the pedagogical aims and approaches implicit in the structure of the skills selected for microteaching were different from those to which pupils were accustomed or even from those to which students were expected to conform in teaching practice

b)/

b) the lecturers' reaction

- while the questionnaire responses revealed views ranging from qualified enthusiasm to mild scepticism about an approach which seemed an additional and perhaps pointless complication of existing training procedures, the overall reaction would seem to illustrate the problems inherent in centre-periphery models of innovation. The drive in this instance came from the researcher associated with the principal and senior lecturers; and the impression which remains is that, although the department as a whole gave its intelligent and loyal support for the duration of the experiment, there was insufficient conviction about the value of microteaching to ensure its development once the thrust from the centre had been withdrawn.

There is no doubt, however, that the experiment aroused a good deal of interest in a number of college departments, and the technique is currently being applied, at varying levels of sophistication and intensity, in the Mathematics, Chemistry, Modern Language and Classics departments and (in a highly organized way) throughout the School of Further Education. In view of this interest and of the positive indications of the History department study, there would seem to be ample justification for carrying out further research. Ideally, such a research programme should have a coherent plan, gradually broadening out from the initial study. The study itself requires replication, involving larger experimental and control groups and extended sampling of classroom teaching (the recording of at least three lessons by each student) at both pre- and post-microteaching stages, to allow for lesson-to-lesson variation in use of skills. The experiment should also be extended to include other subject specialisms and other skills. There is a need, in particular, to examine the range of teacher skills associated/

associated with small group and independent learning, and to explore what modifications (for example, larger groups of pupils, extended practice periods) such skills may imply for the practice of microteaching.

While it is certainly necessary to experiment with a broader range of skills, there is equally work to be done in exploring questioning skills in greater depth. The value, indeed the inevitability, of questioning appears to be taken so much for granted by most people engaged in teacher training that any research worker in this field should perhaps remind himself of Charlotte Mason's comment (in "Home Education") that "oral lessons have their occasional use, and when they are fitly given it is the children who ask the questions." But, assuming the importance of instructing students in questioning skills, it is still a matter for debate (and for further research) whether models of question technique are most aptly presented in terms of a cognitive hierarchy based - as in many research studies - on Bloom's taxonomy. In any future experiment involving questioning skills, it would be interesting to discover whether a model based on teacher purposes (such as that proposed by Hough and Duncan, 1970) might provide a more meaningful framework and result in more effective application. For example, questioning might be broken down into broad purposes such as

- questions to provide feedback on pupils' powers of recall (plus the reinforcement of pupils' knowledge)

- questions to provide feedback on pupils' understanding of material previously presented (plus the reinforcement of pupils' understanding)

- questions to encourage pupils to make fresh steps in comprehension/

comprehension, analysis, synthesis, evaluation (plus feedback to teacher on the capacity of pupils at these levels)

- questions concerned with procedure, relating to the organization and management of the lesson situation.

Any categorization of this kind would need to be based on an extensive analysis of the apparent purpose of questions in recorded lessons.

There is also a need to explore the extent to which students can be trained through microteaching to ask the "right" question in a given context and to phrase it in a clear form. Research studies on questioning have tended to ignore these skills, perhaps because they cannot be assessed in quantitative terms. Yet both the researcher and the supervisor in the Jordanhill experiment noted that phrasing questions in a muddled, ambiguous form (often associated with the repetition of a question in successive forms, without intervening pause for response) was one of the commonest weaknesses among the experimental groups. Research along these lines must accept the need for high inference assessment and must therefore incorporate extensive reliability training for raters.

The separate aspects of the microteaching programme used in the Jordanhill study could, with advantage, be examined in greater depth. A good deal of relevant research evidence on modelling, feedback, supervisor's role, etc. is available, but many of the indications need to be tested within the context of the graduate training year. Lines of further study, suggested by the Jordanhill experiment, include

modelling

- should separate models be prepared of a skill used in a micro-teaching setting and the same skill used in a "normal" classroom? What/

What would be the different functions of such models? At what stage of the cycle should each be viewed?

- what would be the effect of varying the time lapse between viewing the model and the microteaching practice sessions?

number of cycles

- would an extension of the time devoted to the practice of each skill result in more significant gains in subsequent classroom performance?

peer teaching

- no attempt was made in the Jordanhill study to use peers as pupils. A comparison of effectiveness should be made, using pupils and peers, to extend the findings of Levis (1973).

audio v. video feedback

- the possible superiority of video to audio feedback, even in relation to verbal skills such as questioning, should be related to studies of various ways of exploiting video techniques (width of shot, cutting, zooming, screen splitting), to explore the differing aspects of self-concern confrontation, performance confrontation and impact confrontation, as described by Fuller and Manning (1973).

discrimination training

- Wagner (1973) found that students trained (through the presentation of taped examples) to recognize varying instances of pupil-centred teaching behaviour became significantly more pupil-centred in their own teaching than a group which had practised the same techniques but had not been given discrimination training./

training. The findings suggest, not that the practice element in microteaching should be abandoned (since it provides the student with a means of testing the effect on pupils of different types of teaching behaviour), but that the time spent on reviewing skills models could usefully be extended. Wagner's results should be explored further in the context of the graduate training year.

team-work approach

- the favourable reactions of students to group sessions at the feedback stage suggests the value of investigating the performance effects of team-work at the planning and teaching stages of the cycle, along the lines developed at the New University of Ulster.

On a broader basis, two important indications of the Jordanhill study require further examination, particularly since the pattern of research evidence in both areas is inconsistent. First, in terms of coded behaviour (though not in terms of tutors' assessments), the experimental groups showed a significant capacity to transfer at least some of the acquired skills to their subsequent classroom behaviour. In any future research, some means must be found, through extended sampling preferably devised in such a way that the students remain unaware of the particular behaviours being examined, of checking these tentative findings. Second, the evidence (see Table XXXIX) suggesting that students initially weaker in the skills benefited at least as much as students initially superior must be tested by further experiment. The feeling of the College staff, that, if only for practical reasons, microteaching should be used primarily as a remedial technique, is not contradicted by this initial study, but is challenged both by the conclusions of Fuller and Manning (1973) - who felt that "the rich get richer and the poor/

poor get poorer" as a result of self-confrontation - and by the conviction of the Jordanhill History department supervisor, at the end of the 1974 training session, that in a group "too heavily weighted with poor students...perception of weakness and the enthusiasm which is generated by sharing in a lively and effective lesson situation was lost"; a loss which "would seem to present problems if mciroteaching is used as a remedial technique only."

The Jordanhill microteaching experiment was a small pebble thrown into a large pool of practical experience, but, for a time at least, it produced ripples of concern over a wide area: problems of defining, identifying and analysing skills, of relating the use of skills to measures of effectiveness, of assessing the value and limitations of microteaching in relation to other training techniques, of integrating the approaches implicit in microteaching with the training curriculum as a whole, and of gaining acceptance for an innovatory technique within a system which was certainly well tried, if not in an objective sense well tested. Like most exploratory studies, the experiment raised many more questions than it answered; a state of affairs which points to the need for future research over a field much broader than was initially envisaged. The following areas in particular would seem to call for intensive investigation:

(i) the nature and practice of teaching skills

Reference has been made to the hierarchical character of teaching skills. The problem for the trainer of teachers who adopts an analytic approach is to decide the point at which, from the students' point of view, it is best to stop in breaking down each broad competence into its component parts. Borg (1970) suggests that, with traditional training procedures, "teachers are being asked to compose symphonies before they have learnt the notes of the scale"; but he goes/

goes on to remind us of the need to "put the teacher back together again", once the dissection of his performance in terms of technical skills has been completed. The search for the optimum balance between analysis and synthesis should certainly form part of future inquiry into microteaching procedure. It is perhaps a temptation for research in this field to work within a range of narrowly defined sub-skills, since the fewer the behavioural criteria presented to the student at any one time the greater the likelihood of his achieving measurable gains. But it is possible that the student who has spent many hours mastering, say, different aspects of probing may not in his subsequent teaching show himself as effective in question technique as the student who has spent the same length of time practising more extended sequences of questioning behaviour, for example the kind of cycle distinguished by Bellack (1966), involving structuring, soliciting and reacting to pupil responses. The problem is sharpened by the brevity of the graduate training year, giving little opportunity to build gradually from practice of sub-skills to more broadly defined behaviour patterns.

Methods of training students in the appropriate use of skills - "the ability to adapt one's behaviour in a varied way to different goals, situations and responses", which White (1972) considered to be "probably the most important single indication of above-average teaching ability" - would appear to be a logical second stage of any investigation into microteaching techniques. Using a combination of technical skills appropriately is itself a skill, at a higher stage in the hierarchy, where the degree of cognitive processing and decision making lifts the behaviour above the level of straightforward habitual response. The question to be explored is whether training at this level can be adequately given through microteaching. It would seem likely that an intermediate dimension of experience is required/

required, somewhere between the initial microteaching encounter and the uncontrolled classroom situation. Whether such an intermediate training stage can feasibly be incorporated in the graduate training year remains to be determined.

Jordanhill graduates following a Secondary certificate course receive their methods training from their specialist subject departments and, in accordance with this pattern, microteaching experiments have so far been initiated on a departmental basis, so that selected skills have been "subject-related", in the sense of being firmly set in the context of specific subject teaching. Some departments - Modern Languages, for instance - would go further and claim that the teaching of their subject involved certain skills not relevant to other disciplines. There is a clear need for research into the relation of skills to subject matter at the level of secondary school teaching. Investigation of the problem may identify a set of basic skills which are applicable to a very broad content area, while others may be appropriate to a single specialism; but graduates who have been trained to think of themselves as subject specialists may prefer a subject-related approach to the analysis and practice of even the most broadly applicable skills. The issue is important from an organizational as well as a conceptual point of view. Centralised planning of microteaching might effect economies of effort and avoid possible contradictory definitions of a skill, to which a graduate specialising in two main subjects might otherwise be exposed. But centralization would involve considerable changes in the present pattern of training; and there is evidence from microteaching experiments conducted subsequently to the History department study that graduates do indeed like to tackle this kind of work within a framework created by their own specialist department.

If/

If microteaching is to be extended, decisions will need to be taken on the order in which skills should be practised. It would seem logical to follow a hierarchical sequence, as indicated earlier in this section; but whether, at any one level of the hierarchy, there is an optimum order (should the practice of presentation skills, for example, precede or follow the practice of interaction skills?) is a question which requires further investigation.

(ii) relating skills to effectiveness

As indicated in chapter 2, section (b), the links between teaching skills and pupil achievement have not been clearly or comprehensively established. In the restricted period of the graduate training year, selection of the most useful skills (and the level of complexity at which they can most usefully be presented) must be rigorous, and should depend not only on a thorough knowledge of the results of correlational studies already carried out by other researchers but also on local investigations into which patterns of teaching performance appear to work most effectively within the context of curriculum, classroom organization and social background prevailing in the west of Scotland. Studies of this kind (which would seem to be a natural field for college-initiated research and development work) will remain indeterminate until some means can be found of achieving consistent reliability in techniques of high inference assessment. The failure of the initial Jordanhill experiment to demonstrate such reliability in relation to specific skills indicates the pressing need for further research in this area, preferably involving small teams of lecturers in different subject specialisms, working with co-ordinators drawn from the Education or Psychology department.

(ii) strengths and limitations of microteaching

Set/

required, somewhere between the initial microteaching encounter and the uncontrolled classroom situation. Whether such an intermediate training stage can feasibly be incorporated in the graduate training year remains to be determined.

(ii) relating skills to effectiveness

As indicated in chapter 2, section (b), the links between teaching skills and pupil achievement have not been clearly or comprehensively established. In the restricted period of the graduate training year, selection of the most useful skills (and the level of complexity at which they can most usefully be presented) must be rigorous, and should depend not only on a thorough knowledge of the results of correlational studies already carried out by other researchers but also on local investigations into which patterns of teaching performance appear to work most effectively within the context of curriculum, classroom organization and social background prevailing in the west of Scotland. Studies of this kind (which would seem to be a natural field for college-initiated research and development work) will remain indeterminate until some means can be found of achieving consistent reliability in techniques of high inference assessment. The failure of the initial Jordanhill experiment to demonstrate such reliability in relation to specific skills indicates the pressing need for further research in this area, preferably involving small teams of lecturers in different subject specialisms, working with co-ordinators drawn from the Education or Psychology department.

(iii) strengths and limitations of microteaching

Set/

Set against the solid body of research evidence (reviewed in chapter 2, section (c)) that microteaching "works" as a method of training students in new tricks of teaching behaviour, are certain recurrent anxieties (discussed, for example, by Griffiths, 1973; St John-Brooks and Spelman, 1973; Brusling, 1974) concerned with the psychological premises on which the technique is based and its possible inhibiting effects on full pedagogic development. Critics have pointed to the behavioural approach which appears to underlie microteaching, and have stressed the impediment to the development of a personal teaching style inherent in the "processing" of students through a programme of discrete skills; a process in which criteria are selected and standards of performance are set by the trainer, normally with little reference to the views of the students themselves or to individual differences which may call for less standardized training methods. There is concern, too, lest microteaching should act as a conservative force, emphasizing expository class teaching skills, at the expense of more pupil-centred approaches involving the skills of managing small group and independent learning situations. Moreover, as Fuller and Manning (1973) pointed out, training techniques such as microteaching which incorporate an element of self-confrontation are based more on faith and enthusiasm than on clear research evidence about the long-term effects of self-confrontation in changing behaviour or its possibly damaging impact on those students easily subject to stress. Some would criticize microteaching for its prescriptive approach and its lack of emphasis on prior analysis of classroom interaction. Finally, Brusling (1974) suggests that "the need is for a movement away from training instructional interaction skills towards the training of instructional design skills".

In the face of such criticism, Borg (1971), defending the Far West Laboratory's/

Laboratory's Minicourses, argues that the teacher can "combine and adapt these (component) skills to create a unique teaching style best suited to him as a person. Minicourses can no more make a master teacher than can learning the musical scales make a master musician; but they can provide some of the basic tools the teacher needs for professional development." The analogy seems fair and it is arguable that personal teaching style is unlikely to develop fully unless it is founded on an initial competence and confidence in the basic skills. The difficulty is to decide on the means and the moment, within a brief training period, for moving on from the practice of scales to the making of melodies. The kind of transition applied by Ivey and Rollin (1974) (see chapter 2, section (e)(iv)) in the field of microcounselling may well provide an appropriate strategy for graduate training, but it would imply a phased programme of mciroteaching throughout the session. Research into the conceptual basis of such an approach and its practical application might suggest techniques which would alleviate many of the current anxieties about the restrictive nature of microteaching.

(iv) integration of microteaching within the graduate curriculum

The implication of the lines for future research suggested so far is clearly that microteaching must, if its full potential is to be explored, be thought of as an integral part of the curriculum as a whole, not as an extra to be included if time permits; and such integration will inevitably involve change in the established training pattern. On the one hand, there is a need to link the practice of microteaching with the theoretical element in the graduate course, so that students may have a clear grasp of the psychological and pedagogical concepts upon which the technique is based. On the other/

other hand, a bridge must be built between microteaching and school practice, so that movement from one experience to the other involves a steady progression, not an uncertain leap. Some educationists advocate a gradual advance from the controlled microteaching experience to the full-scale classroom encounter; others believe that microteaching comes most appropriately as a reinforcement after the students have been initiated into classroom work. Whatever programme may be preferred within a particular institution - and the decision would need to be tested and refined by prolonged development studies - it is difficult to see how any integrated scheme linking theory with practice can be applied with equal relevance to all students within the training pattern currently in force for graduates at Jordanhill; where half the population tackles practice before theory at each stage of the course, while the other half receives theory before practice.

Setting microteaching in a conceptual framework which embraces both theory and practice will involve the analysis of classroom interactions and the development of teaching models; work which must be carried out at a level appropriate both to the lecturers who plan the training courses and to the students who follow them. Precedents for such integrated approaches have been set at the New University of Ulster and at the University of Stirling, where training in the use of Flanders' interaction analysis categories is developed concurrently with microteaching practice in interaction skills. This pattern of integration is consistent with research evidence. For example, Wragg (1971) reported that students who, after a microteaching practice lesson, received both videotape feedback and an analysis of their performance in terms of Flanders' interaction categories were rated significantly higher by pupils on their re-teach lesson; whereas groups receiving only video feedback or/

or only a F.I.A.C. report on performance showed no significant improvement at the re-teach stage. Wragg's experiment was conducted with only sixteen students, and further studies in this area would help to distinguish the optimum relationship between analysis and practice.

Within such an integrated training structure it will be both possible and necessary to undertake research into the role of microteaching in relation to techniques such as discrimination training, simulation exercises and the stimulation-discussion-action model developed by the Far West Laboratory. Employing microteaching as an isolated "experiment" within a traditional course pattern may result in undue concentration on certain technical skills suited to this particular training formula, to the neglect of other skills - for example, those involving relationships - which may be better tackled by other training techniques. Developing an integrated approach should reduce the danger of merely picking a technique because of its innovatory lustre and allowing it to dictate one's choice of training content.

(v) microteaching and management: the problem of innovation

Ruling with one's pen, as James I appreciated, is a comparatively carefree process. A prescription for integration is simple to write, enormously difficult to implement. Analytic approaches of the kind suggested in the preceding paragraphs involve a good deal of small group work and make considerable demands on the time of both students and lecturers. To accommodate an innovation of this dimension, some other course elements have to be reduced or discarded, and decisions of this order involve a re-assessment of the entire curriculum. In the/

the process, a new pattern of closer interdependence between professional and subject departments may need to be evolved. Research studies cannot presume to dictate such decisions, though they may influence them. But there is a risk that operational research may actually delay large-scale experiment, by raising the kind of awkward questions which produce defensive attitudes. Any attempt from the centre to impose an innovation such as microteaching on departments throughout the College would almost certainly be resisted. Yet sporadic efforts by individual subject departments to use a technique of this kind seem unlikely to develop in any meaningful way, since the theoretical insights from which the approach should draw its conceptual strength are, by tradition, to be found mainly within the Education and Psychology departments. It would be a crude over-simplification of training procedures to suggest that, within a college such as Jordanhill, theory is the province of the "professional" departments, while the subject departments are concerned with practice. Nevertheless, elements of such an over-simplification may well linger in the minds of many graduate students and even of some lecturers. It would therefore seem logical to suggest that the key to any innovatory process which involves a fusion of theory and practice may lie in developing a more integrated faculty structure, perhaps along university lines, where subject specialists work as members of Education departments. Given a community of purpose, the management of large-scale innovation in practical terms should not defeat a College which has very considerable experience of solving the logistic problems involved in handling the training of large numbers of students.

The question remains whether the indications of a tentative initial study provide any justification for pursuing research and development implications to the length of a radical restructuring of the entire graduate/

graduate curriculum. Perhaps the presumption may be excused on two grounds. First, the reactions of the students who took part in the experiment: the very positive response of the Jordanhill group is in accordance with the evidence of studies from a number of countries. It would appear that microteaching provides both intellectual stimulus and psychological support, a blend of qualities not to be lightly rejected. Second, there is a sense in which the study is not yet complete. The purpose was to assess the effectiveness of micro-teaching in the context of the graduate training year. To fulfil this aim, it is not enough to demonstrate that the process can be switched on and made to work. Its true effectiveness can be assessed only when it is wired into a system designed to accommodate it. Future research at Jordanhill could well be concerned with testing the truth of the metaphor.

Jordanhill College History Department

ANALYSIS OF

TEACHING SKILLS

June 1972

PREPARATION

Objectives

1. The teacher was aware of the broad aims to which his immediate objectives should relate.

2. The teacher clearly defined his objectives in terms of knowledge, skills, understanding and attitudes which his pupils should achieve.

3. The teacher chose objectives appropriate in difficulty to the age, intelligence and experience of his pupils.

4. The teacher limited his objectives to what was feasible within the constraints of time and resources available for any particular learning sequence.

Jordanhill College History Department Teaching Skills Analysis

Choice of Content

1. The teacher chose material which was

   a. likely to interest the pupils

   b. at an appropriate conceptual level

   c. related to an important aspect of the period/topic being studies.

   d. related to the objectives, i.e. which enabled the pupils to develop the skills, knowledge and attitudes specified.

Planning of Method

1. The teacher prepared lesson notes which

   a. set out the material in a clear and orderly sequence

   b. planned the wording and timing of key questions

   c. planned how learning aids were to be used as an integral part of the lesson

   d. planned the groups in which the pupils would learn (e.g. class work, group work, individual assignments).

PRESENTATION

Jordanhill College History Department Teaching Skills Analysis

Clarity and Coherence

1. In presenting his material, the teacher spoke audibly and confidently, varying the tone, pitch, etc. of his voice to suit the subject matter and to keep the pupils' attention.

2. The teacher suited his vocabulary, sentence-structure etc. to the level of the class.

3. The teacher clarified the points he was making by using illustrations and examples relevant to the pupils' knowledge and experience.

4. The teacher took care to explain the meaning of any words or phrases with historical connotations differing from their current applications.

5. The teacher clearly explained the historical concepts involved in his material.

6. The teacher developed his material in an orderly sequence.

7. The teacher arranged his material in such a way that the pattern was intelligible to the pupils.

8. The teacher laid stress on the key points of the material.

9. The teacher was alert to the possibility of pupils "losing the thread", and took appropriate steps to clear up any difficulty.

Content

1. The teacher displayed a grasp of content sufficient to ensure

   a. that his exposition was accurate, precise and clear.

   b. that he could deal competently with questions from pupils that were relevant to his theme.

Stimulus

1. The teacher stimulated and maintained interest in his material by

   a. conveying a sense of personal interest in the subject matter.

   b. avoiding monotony in the tone and rhythm of his presentation.

   c. employing an imaginative range of techniques and resources, so as to vary the stimulus.

2. The teacher took constructive advantage of signs of interest on the pupils' part, e.g. questions raised by the pupils, discussion initiated among the pupils, material brought by pupils to the classroom.

Use of Learning Aids

1. The teacher displayed efficiency in organizing and operating the aids that he employed.
(e.g. adequate supply of chalk, duster, drawing pins etc.
technical competence in operating a.v. equipment
positioning of screens, level of blackout etc.)

2. The teacher made sure that all aids (including pictures and diagrams) were clearly seen and/or heard by pupils in all parts of the room.

3. In his blackboard or O.H.P. work, the teacher

- wrote and drew with adequate neatness and clarity.
- laid out his material in an effective manner.
- used the blackboard or O.H.P. to reinforce the key points, unfamiliar names, terms etc., at each stage of his exposition.

4. The teacher gave the pupils appropriate cues before viewing or listening, (e.g. "watch out for this, listen for that"), in order to direct their attention to the most important aspects of the material.

5. The teacher followed up the presentation or study of the material to ensure that important points had been grasped.

6. The teacher involved the pupils, where possible, in the presentation and use of resource material.

Jordanhill College History Department — Teaching Skills Analysis

Varying the Pace & Rhythm

1. Throughout the lesson, the teacher maintained an appropriate balance between his own exposition, teacher-pupil interaction and pupil activity. (In particular, he avoided the lecture approach, in which the pupils spent the whole lesson listening passively to an unbroken flow of talk from the teacher.)

2. In his exposition, the teacher

   a. varied the pace and style of his delivery as appropriate

   b. varied the rate of presenting new facts and ideas, according to their complexity;

   c. employed a range of appropriate resources, to achieve a variety of stimulus.

3. Wherever appropriate, the teacher used questioning and discussion to interact with the pupils and break up the flow of exposition.

4. The teacher took every opportunity of promoting a variety of pupil activities (class, group or individual, as appropriate).

5. The teacher ensured that no one section of the lesson continued to the point where boredom set in, or went on so long that it became necessary to rush or even omit other essential sections.

6. The teacher balanced periods of intensive effort and activity against quieter, more relaxed periods.

Evaluation

1. The teacher employed a variety of measures to assess achievement of the objectives which he had set for the pupils.

   a. by informally checking as the sequence progressed

   b. by formal tests at appropriate intervals

   c. by introducing, where possible, a measure of self-assessment by the pupils

2. The teacher was critically alert to the pupils' reaction in terms of involvement, interest etc.

3. At the end of the sequence, the teacher was prepared to think constructively about amending the course in the light of his own observations and assessments.

Books and Documents

1. In using text-book material, the teacher was careful to

   a. ensure that all references to the text were relevant to the subject matter he was developing

   b. ensure that the material was fully understood by the pupils

   c. make explicit connections with the main theme of his presentation

   d. avoid mechanical reliance on the text book as a convenient alternative to careful teaching and active learning

2. In using historical documents, the teacher was careful to

   a. put the documents in their context

   b. present the documents in an interesting way

   c. make sure that the pupils understood all important words and phrases

   d. where possible, encourage the pupils to draw their own conclusions from the documents

   e. where appropriate, encourage the pupils to consider the value of the documents as evidence (bias, reliability, etc.)

PUPIL INVOLVEMENT

Jordanhill College History Department — Teaching Skills Analysis

Question Technique

1. At all appropriate points, the teacher asked frequent questions.

2. The teacher phrased his questions clearly and concisely.

3. The teacher avoided ambiguous questions (i.e. questions in a form which made it difficult for pupils to decide what was the answer required).

4. The teacher avoided over-use of the "one-word answer" type questions.

5. The teacher refrained from mechanical repetition of the pupils's answers.

6. The teacher distributed his questions so as to involve as many members of the class as possible at different times - (To this end, he used a mixture of simple and more challenging questions, to cater for differing levels of ability.)

7. The teacher gave the pupils sufficient time for thought and formulating of an answer, before resorting to answering the question himself or abandoning the question unanswered. Wherever possible, he insisted on adequate answers, properly phrased.

8. The teacher discouraged undisciplined response to questions (e.g. calling out, wild hand-waving.)

9. The teacher acknowledged good answers with approval and poor answers with tolerance.

10. The teacher made constructive use of pupils' answers, by probing initial responses and re-directing them to other pupils or to general discussion.

11. The teacher varied the level and nature of his questions.

e.g.	Recall	-	Who? When? What happened after? etc.
	Translation	-	What is meant by? What can you see in ---- ? etc.
	Interpretation	-	What were the reasons for? What were the main causes/ consequences? etc.

Pupil Participation

1. The teacher established an atmosphere in which the pupils were willing to participate when called upon to do so.

   He
   - stimulated interest
   - presented material in an open-ended challenging way, involving discoveries to be made, problems to be solved.
   - built up a sense of confidence in the pupils by avoiding an over-critical, repressive attitude
   - instilled in the pupils a sense of being responsible for their own learning.

2. The teacher asked frequent questions and involved as many pupils as possible in answering. (See under Question Techniques).

3. The teacher encouraged the pupils to ask questions provided that they were relevant to the subject in hand.

4. The teacher took every opportunity to encourage purposeful discussion.

5. The teacher varied the modes of learning (class work, group work, individual work) in such a way as to provide for the maximum involvement and activity.

6. The teacher encouraged the pupils to seek information for themselves (rather than giving them all the information himself) and provided the resources necessary to achieve this.

7. Wherever possible, the teacher provided for an element of practical work, some of it involving physical activity.

8. The teacher included an element of written work, both for consolidation of what was to be learnt and as a form of creative expression.

9. The teacher directed the pupils to follow-up activities arising from the topic being studied.

10. The teacher steered clear of aimless and time-consuming "activity for the sake of activity".

Organization

1. The teacher maintained an efficient working environment, in terms of ventilation, lighting, space etc.

2. The teacher showed himself capable of dealing with the logistic problems involved in group and individual work and in practical activities.

3. The teacher organized group and individual work in such a way as to

   a. let everyone know what he had to do and provide him with enough work to keep him busy.

   b. distribute his own attention effectively among the different groups.

   c. allow adequate time for activities, so that pupils could gain some sense of achievement.

4. The teacher displayed an awareness of the physical constraints (environment, resources) under which he had to work, and did not attempt anything over-ambitious.

RELATIONSHIPS

Relationships

1. The teacher established an effective relationship with his pupils by

   a. encouraging their efforts and praising achievement wherever possible.

   b. being willing to listen to what they had to say, and showing an interest.

   c. being ready with help whenever needed.

   d. giving the less able pupils a share of encouragement and attention.

2. The teacher exercised adequate control over the class, so as to maintain an atmosphere of purposeful learning without inhibiting initiative.

APPENDIX B
SECTION

JORDANHILL COLLEGE HISTORY DEPT.

TEACHING SKILLS APPRAISAL GUIDE

(1st version)

Please place a tick in appropriate column against each skill which you are able to assess during the lesson.

STUDENT

VISITED IN

DATE

TUTOR

CLASS

LESSON SUBJECT-MATTER

			Very poor E	Rather weak D	Slightly below average C-	Average C	Slightly above average C+	Very Good B	Outstanding A	NOTES
PLANNING	1.	Objectives	The teacher clearly defined his objectives, of appropriate difficulty and feasible in the time available.							
	2.	Choice of Content	The teacher chose worthwhile material, appropriate to the achievement of his objectives.							
	3.	Planning of Method	The teacher prepared lesson notes setting out clearly the methods by which the pupils were to learn.							
	4.	Choice of Learning Aids	The teacher chose suitable learning aids, showing awareness of the wide range available.							
	5.	Follow-Up	The teacher planned appropriate follow-up work related to his objectives.							
TEACHING SKILLS and BEHAVIOURS	6.	Beginning the sequence	The teacher introduced the material in an appropriate and interesting way.							
	7.	Clarity and Coherence	The teacher presented the material clearly (voice, vocabulary, examples, concepts) and in an intelligible pattern.							
	8.	Content	The teacher displayed an adequate grasp of the subject-matter.							
	9.	Use of Learning Aids	The teacher made effective use of appropriate learning aids (blackboard, AVA, models, etc.)							
	10.	Books and Documents	The teacher used text books and historical documents in a discriminating and purposeful way.							
	11.	Question Technique	The teacher frequently asked questions (phrasing, level, distribution) and handled pupils' answers constructively.							
	12.	Pacing	The teacher effectively controlled the pace and rhythm of of his presentation.							
	13.	Pupil Participation	The teacher provided for a variety of worthwhile pupil activity							
	14.	Evaluation	The teacher adequately assessed the achievement of the objectives set for the pupils.							
	15.	Ending the Sequence	The teacher brought his sequence to a satisfactory end (revision, consolidation, involvement, follow-up)							
	16.	Organisation	The teacher maintained the conditions for efficient learning as far as lay within his power.							
	17.	Stimulus	The teacher presented his material in a stimulating and imaginative way.							
	18.	Relationships	The teacher established a good relationship with the class and controlled them effectively.							

General Comments

Total Score

No. of skills assessed

Mean Score

Overall Teaching Assessment

Signed

JORDANHIL COLLEGE HISTORY DEPT.

TEACHING SKILLS APPRAISAL GUIDE

Please tick your assessment of
- each general teaching
- each section of specific skill

STUDENT SECTION

VISITED IN

DATE

TUTOR

CLASS

LESSON SUBJECT-MATTER

GENERAL TEACHING SKILLS	Weak	Well below average	Slightly below average	Average	Slightly above average	Well above average	Very good	COMMENTS
	E	D	C-	C	C+	B	A	
1. PREPARATION								
- Clearly defined objectives, appropriate to class								
- Choice of worthwhile content								
- Planning of method & resources								
2. PRESENTATION								
- Clear, coherent exposition								
- Adequate grasp of content								
- Stimulus (to capture & maintain interest)								
- Effective use of learning aids								
- Varied pacing and rhythm								
- Achievement of objectives assessed								
3. PUPIL INVOLVEMENT								
- Efficient question technique								
- Encouraged variety of worthwhile pupil activity								
- Pupil activities well organised								
4. RELATIONSHIPS								
- Established good learning atmosphere								
- Maintained efficient control								
- Developed friendly rapport								

SPECIFIC SKILLS FOR MICRO TEACHING	E	D	C-	C	C+	B	A	
QUESTIONING								
1. Avoiding over-use of 'one-word answer' type questions								1.
2. Avoiding over-use of 'yes/no' type questions								2.
3. Optimum use of questions demanding interpretation, judgement, etc.								3.
4. Giving pupils time to think out adequate answers								4.
5. Use of prompting, to encourage well formulated answers								5.
6. Constructive use of answers, through probing (for clarification, extension of response)								6.
7. Development of discussion, through re-direction of initial answers for further comment								7.
								8.
								9.

Overall Teaching Assessment [ ]

Overall Assessment of Micro Teaching Skill [ ]

Signed .....

Jordanhill College History Department

Appraisal of Teaching Skills

(notes for lecturers)

A. Introductory Notes

1. The History department has agreed that in session 1972-73 all students should be assessed in terms of the Teaching Skills Appraisal Guide (a sample copy is attached).

The General Teaching Skills will apply to all lessons throughout the year. In each half term, students will be asked to pay special attention to one of the four "basic teaching skills" which have been selected for special study under micro-teaching conditions. The allocation will be

Term 1	-	1st half	-	Clarity and coherence of exposition
	-	2nd half	-	Question technique
Term 2	-	1st half	-	Varying the pace and rhythm
	-	2nd half	-	- ditto -
Term 3	-	1st half	-	Using historical documents
	-	2nd half	-	No special skill

2. The Appraisal Guide summarizes the Analysis of Teaching Skills (copy attached). The sub-headings on the left-hand page of the Guide correspond with each skill in the Analysis and are simply aide-memoires to assist lecturers in recalling the detailed breakdown.

It goes without saying that the better acquainted lecturers are with the Analysis the greater will be the chances of achieving consistency and reliability in assessing students. The greater the degree of reliability, the greater the validity of our research findings.

3. It will help to achieve reliability if lecturers keep in mind the following points in completing the Appraisal Guide:

a. The four main assessments (Preparation, Presentation, Pupil Involvement, Relationships) and the "specific skill" assessment should be based on the points set out in the Analysis. Other considerations should not be taken into account. To achieve consistency, there must be some agreed basis for measurement. The Analysis may well have its imperfections, but it should be accepted as a basis during 1972 - 73.

b. There can be no consistency unless lecturers agree on a concept of the "average student". For the purpose of scoring the Appraisal Guide under its different headings, lecturers should take "average" to mean "average performance in terms of all the students whom you have assessed". Do not apply different standards of judgment as between

- 1st term and 3rd term students
- Honours and Ordinary graduates
- M.S.I and M.S.II students
- male and female students etc.

If lecturers feel inclined, they can safely give these distinctions some weight in awarding the "Overall Teaching Assessment", since the research project will not be taking this into account. It stands to reason, however, that since our research is concerned with comparing the progress of different groups of students, our detailed assessments must be based on standards which are common to all.

3. c. It may help lecturers to keep in mind the following rough distribution of student ability:

E	D	C-	C	C+	B	A
10%	15%	15%	20%	15%	15%	10%

4. It is impossible to measure teaching performance in purely quantitative terms, and a subjective element therefore necessarily enters into assessment. This raises the difficult question "weighting" - different lecturers will attach more or less importance to different skills. Since, for the sake of simplicity, each of the main assessments in the Appraisal Guide summarizes a number of skills, "weighting" may well reduce consistency of judgments. This cannot be entirely avoided, but the detailed skills analyses have deliberately been constructed as objectively as possible, and it should help to reduce discrepancies of judgment if lecturers can base their assessments closely on these analyses.

5. The Appraisal Guide is meant to be sufficiently comprehensive to apply to all student lessons. Obviously, some of the skills will not be relevant to all lessons. The "use of learning aids", for example, will not always be appropriate. In making their assessments, lecturers should therefore distinguish between

   a. skills which are not used, simply because they are not relevant to that particular lesson. (In such cases, the absence of the skill should obviously not detract from the over-all assessment.)

   b. skills which the student omits to practise (or practises inadequately) in circumstances where the skill is relevant to the achieving of the the stated objectives. (In such cases, the student has committed a "sin of omission", and the assessment should reflect this.)

6. In using the Appraisal Guide, lecturers are asked to take a "double view" of student lessons:

   - a broad, over-all view, leading to the four main assessments;
   - a more analytical, close-up view of the specific skill selected for micro-teaching.

   Taking a balanced view of the wood and at the same time making a detailed study of a particular patch of trees within it is a skill that needs to be practised. Each lecturer will, of course, devise the method of assessment that suits him best. It is hoped that the Appraisal Guide will prove to be an instrument that is accurate without being too unwieldy.

B. The four "basic skills"

I. Clarity and coherence of exposition

General Notes

1. In scoring Clarity and Coherence, choice of subject-matter and objectives must be taken for granted. If you feel that these are badly chosen, you will already have faulted the student under "Preparation". Do not fault him again under this heading

2. Varying the rate of new facts and ideas according to their complexity can affect clarity and coherence, but you will already have taken this skill into account under "varied pacing and rhythm".

Analysis

1. Audible, confident speech

Inaudibility, monotony, diffidence of speech should be faulted under this heading, but not a "rough accent", unless it limits intelligibility; nor should slang or colloquialisms be faulted under this heading (though you might wish to consider them under 2.)

2. Suiting vocabulary, etc. to level of class

Faults under this heading often become apparent through the pupils' reactions, as the lesson progresses. Do not forget to mark positively if the student is clearly making a deliberate effort to match the level of the class.

3. Use of illustrations and examples

"Illustrations" is used here in a verbal sense. It does not refer to visual aids.

Remarks under A (5) apply. Some subjects may not call for many illustrations or examples, others will. Mark the student's performance accordingly.

4. & 5 Explanation of historical words, phrases, concepts

Again, remarks under A (5) apply. If the use of these skills is not relevant to a particular lesson, simply leave this section of the appraisal sheet blank, or write "not applicable". However, it is likely that most pieces of exposition will involve these skills.

6. Material developed in an orderly sequence

"Orderly sequence" is what is meant by coherence.

7. Pattern of material intelligible to the pupils

This is a development of 6. A lesson taught at an inappropriately advanced level might have an orderly sequence which was nevertheless too subtle for the pupils to grasp.

8. Stressing the key points of the material

"Underlining" or "sign-posting" the key points of the exposition may be done in a number of ways: e.g. verbally, by use of blackboard or overhead projector, by recapitulation (perhaps involving questions to pupils).

9. Alertness to possibility of "losing the thread" and clearing up of difficulties

   Note that there are two aspects of this skill:

   a. the teacher should check (usually by questioning) that difficult concepts in his exposition are being followed by the class and should be on the watch throughout for any visual signs of bewilderment;

   b. the teacher should clear up difficulties as he goes along, by working over the ground more slowly and using questions to confirm understanding.

   Once again, A (5) applies. If the pupils are obviously following an exposition which contains no particular difficulties, the student should not be faulted if he does not stop to check up; but he should be faulted for "sins of omission" if the material is difficult and/or the class is not quick on the uptake.

II. Question Technique

General Notes

1. As with Clarity and Coherence, do not let choice of subject matter and objectives enter into your assessment of this skill. You are concerned here simply with the technique of questioning.

2. Often, the effectiveness of a lesson depends not upon asking a great many questions, but on choosing the right questions to ask at the right moment. The analysis of Question Technique does not cover this, and you should not let it affect your assessment. Your judgment of the student's success in choosing the right questions will be reflected in your scoring under "appropriate objectives", choice of content, "planning of method" and "achievement of objectives".

Analysis

1. Asking questions at all appropriate points

   Obviously, lecturers must be their own judges of when questions are appropriate. Discretion under A(5) applies.

2. Clear, concise phrasing of questions

   This speaks for itself.

3. Avoiding ambiguous questions

   This links with 2. Vague, over-generalized questions make it difficult for the pupils to know what answer is expected of them.

4. Avoiding over-use of "one-word-answer" questions

   Obviously, there may be occasions when the one-word-answer question is appropriate; but over-use - particularly when the teacher leads with a complete statement and pauses expectantly to allow the pupils to provide the "missing word" - can limit pupils' thinking and inhibit their verbal expression.

5. Avoiding mechanical repetition of pupils' answers

Again, repetition of answers may sometimes be useful, for reinforcement or to make sure that everyone has heard; but the teacher who mechanically repeats every answer is implicitly encouraging pupils to content themselves with mumbled replies, while the other pupils listen not to them but to the teacher.

This skill implies that one aim of questioning is to make the pupils active, to bring them into the foreground. Pupils should be encouraged to answer so that other pupils can hear. The degree of "teacher-dominance" will thereby be reduced.

6. Distribution of questions to involve as many pupils as possible

Lack of this skill is common among inexperienced teachers. The situation in which the majority of the class remains passive and excluded while answers are accepted from the bright and willing minority, will be well-known to lecturers.

A proportion of open-ended questions (i.e. questions to which there is more than one possible reply) will assist distribution.

7. Sufficient time given for thought and formulating of answers. Insistence on adequate phrasing of answers.

Again, lecturers will recognize the type of young teacher who is so eager to pour out the maximum amount of information that he merely makes a gesture towards questioning. If the right answer is not obtained immediately, he either answers the question himself or abandons the question altogether. Pupils soon come to recognize this trait in a teacher and are discouraged from making the positive effort of thinking that questioning should involve. Properly phrased replies are part of the process of precise thinking.

8. & 9. Discouraging undisciplined response and acknowledgment of answers.

These skills are aspects of good relationships between teacher and class. Intolerant reaction to poor answers can result in embarrassment, resent- and a consequent inglorious muteness on the pupils' part.

10. Making constructive use of pupils answers (prompting, probing, re-direction)

Prompting involves giving cues to a pupil whose first response is incomplete or inaccurate, thus leading him on to a more adequate reply.

Probing encourages a pupil to explain his first answer more clearly or to go beyond his first response by thinking about and commenting on its implications.

Re-direction of a response (to another individual pupil or to the class as a whole) is a way of opening a subject up and involving the maximum number of pupils in an active role. Too often, questioning means spending time on one pupil while the others switch off. Re-direction can encourage a class to feel involved in a pupil's reply.

11. Varying the level and nature of questions (recall, translation, interpretation)

The appropriate balance among types of questions will of course vary according to the nature of the subject-matter and the ability and experience of the class. The skill assumes, however, that teachers should try to extend the range of their questions as widely as possible. There is a tendency among inexperienced teachers to concentrate on recall questions, thus restricting pupils' thinking to a low cognitive level.

III Varying the pace and rhythm

General notes

1. This is a difficult skill to analyse and assess. It is concerned with the flow and balance of the lesson over-all, with the total effect, the sum of the parts.

2. The skill certainly cannot be assessed in quantitative terms. Inevitably, an element of subjective judgment will enter into assessment. Nevertheless, most experienced observers will agree on the importance of correct pacing and will be quick to detect its presence or absence in a lesson. It is, moreover, a skill which many teachers in training need to practise.

3. The skill is distinct from clarity and coherence of exposition. It lies in a proper blend of exposition, interaction and pupil activity.

4. As in music, we distinguish between pace and rhythm. A 40-minute lecture could be varied in pace, but it would remain a lecture. Varying the rhythm involves changing the "time signature" of teaching and learning, normally by ringing the changes on exposition, interaction and activity.

5. To assess this skill, particularly under point 5. of the Analysis, it may be necessary to refer to the student's lesson plan.

Analysis

1. Maintaining a balance between exposition, interaction, activity. Avoiding lecture approach.

   The "lecture approach" is easy to recognize; but merely interspersing the exposition with odd questions at regular intervals can produce an unvarying, set pace, almost as monotonous as a straight lecture. Habitually devoting the last five minutes of a lesson to written work can also produce a kind of rigidity. The question to consider in assessing this skill is "to what extent is the teacher using a balance of exposition, questioning, discussion, activity, etc., to produce a variety of pace and rhythm which will provide the pupils with fresh stimulus?"

   Note that the balance must be "appropriate"; i.e. appropriate to the subject-matter and to the age and stage of the class.

2a. Varied pace and style of delivery

Note here the extent to which the teacher varies pace and pitch of delivery, gesture, movement, etc. as appropriate, to suit his different roles

- as communicator of hard-core information
- as story-teller
- as questioner
- as prompter of discussion
- as manager of activities, etc.

2b. Varied rate of presenting new facts and ideas, according to complexity

This skill is concerned with optimum pacing. Obviously, the pace at which difficult concepts are presented will be slower than the pace appropriate to a series of simple facts.

2c. Range of resources to achieve variety of stimulus

"Resources" will include

- the use of aids such as pictures, slides, diagrams, transparencies, records, tapes, models, exhibits;
- the use of documents and other source material;
- demonstrations, to explain concepts (e.g. the technique of a ballista can be demonstrated by ruler and pellet);
- effective use of blackboard.

3. Interaction through questioning and discussion

This part of the skill implies that questioning should be more than a routine, "throw out the odd question as you go along" affair. A sequence of intensive questioning can break up the flow of exposition. The teacher should be really interested in involving the pupils and in getting them to think. It is this genuine interaction which helps to vary the pace.

Pupil-initiated questions can change the pace again. Young teachers often hesitate to encourage them, perhaps suspicious that pupils will take advantage and waste time.

Discussion can be stimulated by the teacher through open-ended questioning and re-direction of initial responses.

4. Promoting a variety of pupil activites

In assessing this skill, it is worth keeping in mind the many forms of activity which can ring the changes on the usual piece of writing at the end of the lesson. For example -

- Discussion in pairs or groups, leading to prepared statements
- Debate

/-

4. (contd.)

- Role-playing and simulation
- Practical activities (drawing, modelling, etc.)
- Demonstration by pupils, and assistance with presentation of aids, blackboard work, etc.

N.B. In assessing 3. and 4., watch for positive merit on the one hand and "sins of omission" on the other.

5. Avoiding boredom, rush, omissions

Examination of the lesson-plan will be useful in assessing the student's ability at this skill - the skill of controlling the content so as to avoid being trapped in a situation which, in terms of pacing, gets out of control. We need to know what the teacher intended to do when he started the lesson.

6. Balancing intensive effort against more relaxed periods

This assessment will be made by way of a summing up of the flow of the lesson as a whole. It would be possible for a teacher to ring all the changes from exposition to activity without achieving this kind of "peaks and valleys", "light and shade" quality. This is a subtle, sensitive skill, at the heart of pacing and rhythm.

IV. Using Historical Documents

General Notes

1. "Using Historical Documents" should be thought of as a Presentation skill, although it is not separately listed under that heading in the Appraisal Guide. In the first half of the summer term, students will be asked to include this skill in all lessons taught for assessment by lecturers, and you should score it in detail as the "specific skill for micro-teaching". If documents are introduced into lessons observed by you in the first and second terms, you should assess the skill in more general terms, as part of the student's performance in Preparation and Presentation.

2. Documents can be used simply as an introduction to a more general topic or as a peg on which to hang a series of recall questions. Such a use may be legitimate, but it is not what we are concerned with in this context. For the purpose of training students in this skill, the document should form the core of the lesson sequence, and the teaching and learning should be based directly upon it.

Analysis

1. Putting documents in their context

   An obvious aspect of the skill, but one that has been neglected by many of the students practising under micro-teaching conditions.

2. Interesting presentation

   Documents can make a dull lesson if they are presented unimaginatively. Interesting presentation techniques might include verbal scene painting; supporting visual or aural aids; and (if the document is visually attractive, e.g. a poster, play-bill, etc.) effective visual presentation of the document itself, perhaps through overhead or slide projection.

3. Understanding of important words and phrases

   This is a "translation" skill - making sure that the pupils can accurately render the meaning of the document in their own terms.

4. Encouraging pupils to draw their own conclusions.

   This is an "interpretation" skill. The range of conclusions that pupils might be expected to draw would include

   - the character of the times that the document reflects
   - the situation which prompted the writing of the document
   - the character of the author
   - the immediate reason for writing the document
   - the immediate effect upon the readers
   - the long-term influence of the document.

5. Value of the documents as evidence (bias, reliability, etc.)

   Not all documents lend themselves to this treatment, so that A (5) may sometimes apply. Assessing the value of documents as evidence can be a testing exercise for pupils (involving consideration of internal and external evidence and perhaps comparison of two differing reports of the same events), so that lecturers should take into account the ability of the class as well as the nature of the documents, in determining when it is "appropriate" to exercise this aspect of the skill.

APPENDIX E

Jordanhill College History Dept. Teaching Skills Analysis

QUESTION TECHNIQUE

STUDENT

SECTION

DATE

TUTOR

SUBJECT-MATTER

NUMBER of PUPILS DURATION TEACH/RE-TEACH

	Very poor 1	Rather weak 2	Slightly below average 3	Average 4	Slightly above average 5	Very good 6	Outstanding 7
1. The teacher phrased his questions clearly and concisely.							
2. The teacher distributed his questions so as to involve as many members of the class as possible at different times.							
3. The teacher gave the pupils sufficient time for thought and formulating of an answer, before answering the question himself or abandoning the question unanswered.							
4. Wherever possible, he insisted on adequate answers, properly phrased.							
5. The teacher refrained from mechanical repetition of the pupils' answers.							
6. The teacher acknowledged good answers with approval and poor answers with tolerance.							
7. The teacher made constructive use of pupils' answers, by probing initial responses and re-directing them to pupils or to general discussion.							
8. The teacher varied the level and nature of his questions (e.g. Recall, Translation, Interpretation).							
Notes & Comments							

Total Score

Mean Score

% age Score

(sgd.)

APPENDIX F

Jordanhill College History Department

History teaching in Secondary Schools

Four basic questioning skills

(notes for students)

Session 1972 - 73

1. Introduction

The skills of teaching can be listed under a number of headings, for example -

a. Planning skills

- selecting and defining learning objectives appropriate to the class which you have to teach
- choosing worthwhile content, related to the objectives
- selecting suitable teaching and learning methods and planning the effective use of resources

b. Presentation skills

- clear, coherent exposition of material
- adequate grasp of content by the teacher
- stimulating and maintaining interest
- effective use of learning aids
- varying the pace and rhythm of the lesson overall
- checking to ensure that objectives are achieved

c. Skills to achieve pupil involvement

- efficient question technique
- encouraging a variety of worthwhile pupil activity (e.g. discussion, written work, practical activities)
- organizing group and individual work efficiently

d. Relationship skills

- establishing a purposeful learning atmosphere
- maintaining adequate control
- developing a sense of trust and confidence among pupils in their relationships with the teacher.

Listing the skills of teaching in this way implies that teaching is not simply a mystery, a gift that you are born with, but a craft that can be to some extent analysed and practised. You will recognise that many of the skills listed above are common to all teaching situations. Others may apply only to certain types of lesson. In the classroom, an experienced teacher selects/-

selects almost unconsciously, from the battery of skills at his disposal, those that are appropriate to a particular set of objectives, to a specific choice of subject-matter and to the age and stage of a particular class. In the same way, an experienced driver reacts almost automatically to each different traffic situation, choosing the appropriate gear, applying the correct degree of braking or acceleration; but just as a driver has to begin by consciously practising each driving skill (gear-changing, steering and so on), so must a teacher begin by conscious analysis and practice of each aspect of his craft. Choosing the correct blend of skills, making an appropriate synthesis - this can be called the "art" of teaching, and it takes time to develop.

In your teaching practice this session, we are asking you to pay special attention to four basic questioning skills. We have chosen to concentrate on aspects of question technique for three reasons:

(i) Questioning is one of the most common forms of teacher-pupil interaction, forming some part of almost every lesson;

(ii) Questioning is a skill that can be analysed and consciously practised;

(iii) Ineffective question technique is a common weakness among inexperienced teachers.

In the following pages, you will find an analysis of the four questioning skills that have been selected for study. Read the analysis carefully and refer back to it when you come to do your teaching practice. Your tutors will be assessing your use of these skills when they visit you in schools to observe your teaching.

Finally, remember that the ultimate criterion of teacher effectiveness is pupil performance. Teaching skills are not an end in themselves, but a means to the end of better learning. In your teaching, you must therefore be sensitively aware of pupil reaction (as well as taking deliberate steps to assess pupils' achievements of the objectives you have set for them), and be prepared to adapt your basic skills to the particular needs and capacities of your class.

Students who are interested to explore the subject of Questioning Skills in more depth should read an article by Gall, M.D., "The Use of Questions in Teaching", published in vol. 40 no. 5 (December 1970) of the Review of Educational Research. A copy of this article is available in the Library, on application at the inquiry desk.

Four basic questioning skills

1. Avoiding over-use of the "one-word answer", "yes/no" type of question.

2. Varying the level and nature of questions, with particular attention to increasing the proportion of "higher order" questions, demanding interpretation, judgement, etc.

3. Giving pupils time for thought in formulating answers and, through prompting, encouraging adequately phrased responses.

4. Making constructive use of initial answers by

   a) probing, for clarification and extension of response,

   b) re-directing initial answers for comment and discussion.

General comments

Asking questions has for so long been accepted as a basic characteristic of teachers that a trainee may dutifully sow his lessons with questions, without being too sure of what he wants to grow from them. Questioning may be valuable as

a. a form of feedback, enabling the teacher to discover quickly what the pupils already know, how much they have remembered or understood of material previously taught;

b. a means of stimulating thought among the pupils. This applies particularly to so-called "higher-order" questions, i.e. questions which involve more than recall or simple description, demanding a process of reasoning and judgement;

c. a stimulus to verbal expression, giving the pupils an active role, providing (through open-ended questions and through re-direction of initial responses) a social element, reducing the dominance of the teacher.

Too often, questioning is conducted in a perfunctory manner. The teacher, determined to "get through" a prepared amount of material by the end of the lesson, is interested only in the answer which enables him to move quickly to the next stage of his exposition. Under certain circumstances, this may be a necessary strategy, but it ignores the value of questioning as a stimulus to thought, when the teacher, deeply interested in what the pupils have to say, is prepared to follow in whatever direction the answers take him, even if this means straying from a pre-arranged route.

The four skills listed above do not by any means cover the whole technique of questioning. The teacher must also pay attention to

- the clarity with which he phrases his questions
- the number of questions that he asks

/-

- the appropriateness of his questions (key questions should be written into the lesson plan)

- the distribution of his questions, to involve the maximum number of pupils

- the manner in which he acknowledges response (approval of good answers and tolerance of poor ones).

However, the four selected skills lend themselves well to deliberate practice. As you can see, the first two skills are concerned with the style and level of posing questions, and the second two with the way in which the teacher should deal with answers.

Specific analysis

1. "One-word answer", "yes/no" questions

Rapid-fire questioning, demanding one-word answers from a wide scatter of pupils, is a common practice among teachers. It can be useful as a quick means of sampling the information level of the class, and some would justify it as a stimulus to pupil alertness. However, questions phrased in a form that demands only a one-word answer can limit the range of pupils' thinking and inhibit their verbal expression. This is particularly noticeable when the teacher leads with a statement and pauses expectantly for the pupils to supply the missing word. ("The next King of Scotland was James the .....?")

Questions which demand nothing more from the pupil than a "yes" or "no" may also be occasionally justified, but their over-use should be avoided. Feedback from "yes/no" questions is unreliable, since a pupil always has a 50% chance of guessing a correct answer. Questions in this form give no practice in verbal expression. Moreover, even if the teacher intends his "yes/no" type of question to stimulate thought, he can obtain no sure evidence that thinking has taken place. For example, the answer "yes" to the question "Do you think that James IV was an able king?" may be the result of considered analysis and judgement on the pupil's part; or it may simply be a hopeful attempt to provide the answer that the teacher wants, with the minimum of mental effort. Re-phrasing the question in the form, "Why do many people rank James IV as an able king?" will be likely to produce replies that are more carefully thought out and more fully expressed; and (since the question is open-ended, i.e. it provides for more than one correct answer) more pupils can be involved and discussion can be developed.

2. Varying levels of questions, with particular emphasis on questions demanding interpretation, judgement, etc.

A good many attempts have been made at classifying questions according to the nature and level of the thought processes that they are meant to provoke. A very simple classification could be

a) questions demanding <u>recall</u> of factual knowledge

b) questions/-

b) questions to test straightforward comprehension of material and the ability to translate or paraphrase the material

c) questions that call for an effort of interpretation, analysis, synthesis or evaluation on the part of the pupil.

Each level of questioning can be associated with certain characteristic "question stems", for example

RECALL	Who? What was the name of? When? What happened next? etc.
COMPREHENSION	What is meant by? What is another word for? How would you put this in your own words? Can you describe what you see in this picture? etc.
THOUGHT	Why? What were the reasons for? Can you account for? What is your opinion of? How would you judge? What comparisons can you make between? etc.

Research studies over the past sixty years have consistently shown that about 60% of teachers' questions are concerned with recall of facts; 20% require pupils to exercise a higher level of thinking; and 20% are procedural (e.g. "Who doesn't have a pencil?" "How many of you have finished the exercise?") It is interesting to note that the proportions have not changed over the years, in spite of the increasing emphasis put upon the importance of getting pupils to think for themselves and of involving them actively in their own learning. Inexperienced teachers in particular tend to ask a high proportion of recall questions, perhaps because they become discouraged in their first few lessons by the poor response to "higher order" questions couched in a form that is too complex for the pupils to follow. Obviously, you need to test the level of information and basic comprehension to ensure that your "thought" questions are appropriate to the pupils' knowledge and understanding. A series of straightforward recall questions can also be a good way of "warming the class up" and of establishing a confident basis for further effort. You should therefore plan your questions at varying levels, but try to increase the proportion of questions that will stimulate thought among the pupils, encourage longer, more carefully reasoned replies and give the class a sense of sharing responsibility in its own learning.

There may be occasions when you are not sure whether a pupil's reply to a "higher order" question is the result of independent thought or merely recall of material previously heard or read. You can test this by probing the initial response, in the manner described in 4. below.

3. Giving pupils time for thought and using prompting to encourage adequate responses.

Pupils soon come to recognize the type of young teacher who is so eager to pour out a mass of information that he merely makes a gesture towards questioning. If the right answer is not obtained immediately, he either answers the question himself or abandons it completely. This soon discourages the pupils from making the positive effort of thinking that questioning should involve. They are content to sit back and let the teacher do the work.

Whenever you pose questions that demand an effort of thought, you should pause to give the pupils sufficient time to formulate an adequate response. Inexperienced teachers may fear that pausing in this way may simply produce lethargy or distraction among the pupils. You can meet this danger by cueing the class so that they know what is expected of them, e.g. "Now, think carefully about this before you reply', "Take your time and give me a complete answer".

The initial response to your question may be inadequate in a number of ways. It can be wholly or partly inaccurate, it can be incomplete, or it can be poorly expressed in a confused manner. If you dismiss the inadequate response and turn immediately to someone else, you are in effect discouraging the first pupil from making any further effort. If you accept a poor reply without comment, you are effectively lowering the pupil's standards and your own. You can avoid both these dangers by prompting the pupil to improve on his first response. Prompting consists in providing a series of stepping-stones for the pupil who cannot achieve an adequate response in one leap. You can re-phrase your question in a simpler form. You can give the pupil hints that will help him, through a process of recall or association, to find his way to an accurate answer. You can ask him to add to his first attempt or to express it in a clearer form. This procedure requires patience on your part and it may slow down the rate of your teaching (not necessarily a bad thing), but it will be time well spent from the pupils' point of view, standing them in good stead when, at a later stage, they are required to express their ideas clearly and accurately on paper. You will also be reminding your class that you require consistent standards of clear expression. Well phrased answers are part of the process of active, precise thinking and the apparently inarticulate pupil may well respond to training in good habits of speech.

4. Constructive use of initial answers

a. Probing

This is the technique of asking a pupil supplementary questions to extend his first response. Probing is employed, at a higher level than prompting, when the first response is acceptable but not up to the standard which the teacher wishes the pupil to attain.

Probing/-

4. a./-

Probing questions can seek additional information, if the first answer is incomplete; clarification, if the first answer is obscure; or justification ("What are your reasons?" "Why do you think?" etc.), if the teacher feels that the pupil has guessed at a reply without giving sufficient thought to it. Through probing, the teacher can control the extent and difficulty of the pupils' learning tasks, in order to produce the maximum effort of thinking and expression.

b. Re-direction

This is the technique of referring an initial response to other pupils for comment and discussion. (e.g. "Would you agree with that opinion?" "Can you add any other reasons?" etc.) Redirection is a way of opening a subject up and involving the maximum number of pupils in an active role. Too often, questioning means spending time on one pupil while the others switch off. Re-direction can encourage a class to feel involved in a pupil's reply and to listen and respond to each other as well as to the teacher.

You will see that the practice of these four questioning skills should produce a situation in which the teacher talks less and the pupils talk more; a style of learning which brings the pupils more into the foreground, reducing the dominance of the teacher. Research evidence indicates on the one hand that teacher dominance is relatively ineffective, once you move beyond the lower cognitive levels of learning; and on the other hand that pupils achieve more if they are actively involved and feel that they have a responsible share in their own learning.

In practice,

It's very easy to slip into the habit of asking quick questions [illegible] simple, one-word answers, including the answers 'yes' or '[illegible]'

APPENDIX G

Model Script
Asking Questions

Rehearse and Record

8 November '72 (teaching sequences)
0930 - 1230

11 December '72 (presenter's links)
1330 - 1630

Teacher's Voice

(asking questions)

fade teacher's voice

Presenter You'll remember that your Notes on Questioning Skills distinguished three broad purposes in questioning. These were

- questioning for feedback, when you want to discover how much your pupils already know, what they've remembered, how well they've understood;

- questioning to stimulate thought and reasoning, to encourage fresh insights and deeper understanding;

- questioning to develop verbal expression, to involve the pupils actively in response.

It's very easy to slip into the habit of asking quick questions that call for simple, one-word answers, including the answers 'yes' or 'no'. Here's an example of what I mean.

INSERT 1 (Take II)

Teacher/

APPENDIX G (Contd.)

Teacher asking a sequence of one-word answer questions.

Presenter You may feel that there was nothing wrong about that. At least the teacher was obtaining a kind of feedback from his pupils. But the feedback from the yes/no questions was unreliable, since the teacher had no means of telling whether the pupils had really thought about their answers, or whether they were just making hopeful guesses. Moreover, questions like "What was the name of the King?" don't stimulate reasoning or judgement - they simply call for a quick act of memory. Finally, this style of questioning does nothing to encourage verbal expression. The teacher was doing at least ninety per cent of the talking, and probably ninety per cent of the work.

Of course, one-word answer questions have their place, usually at the level of recalling information. They can provide the teacher with a quick means of testing the knowledge of the class. But even at this level, you can increase pupil response by altering the form of your question. For example -

---

INSERT 2 (Take II)

Teacher asks yes/no question.

Presenter Any pupil answering that question had a fifty per cent chance of being right, even if he was only guessing. So try again -

---

INSERT 3 (Take II)

Teacher asks the same question in one-word answer form.

Presenter The single-word answer to that question demanded an act of memory but no effort of verbal expression. But consider this -

---

INSERT 4 (Take II)

Teacher asks the same question in 'complete response' form.

Presenter Note that questions phrased in that form enable the teacher to train his pupils into the habit of answering in complete phrases or sentences, and that's a useful skill. So make a conscious effort in your teaching to avoid the over-use of one-word answer questions. They have their uses, but many teachers ask far too many of them.

A good teacher controls the demands that he makes on his pupils by varying the level of his questions. For simplicity, we can classify questions into three categories:

- questions demanding recall of knowledge
- questions involving simple comprehension of material
- questions that call for a reasoned analysis of information, perhaps leading on to a fresh synthesis or an original judgement.

Many teachers rarely move beyond the first level in their questioning. In your own teaching, you should make an effort to increase the proportion of questions that stimulate thought and involve your pupils in active response and discussion. Questions of this kind are characterised by words such as

- Why
- Explain
- Justify
- Compare.

Listen/

Listen now to an experienced teacher varying the level of his questions, blending knowledge, comprehension and analysis; and note how the length and complexity of the pupils' responses vary according to the level of the teacher's questions.

In his lesson, the teacher is dealing with the Battle of Stirling Bridge; but instead of simply presenting the pupils with a list of facts, he begins by getting them to think for themselves about the tactics of mediaeval warfare; then he takes them through the stages of the battle, asking them to analyse and predict, on the basis of the data that he has built up; and finally he asks them to apply their knowledge, in order to work out an alternative strategy that might have given victory to the English.

INSERT 5A

Lesson extract I

Presenter A mixture of questioning levels there: knowledge, comprehension and analysis. The teacher then set up a hypothetical situation.

INSERT 5B

Lesson extract II

Presenter The teacher went on to establish the role of the mounted knights; and then, with the aid of the magnetic board, he showed how the Scots and English deployed their forces. We'll rejoin the lesson at the point where the/

the Scots are about to attack the English at the Bridge.

INSERT 5C

Lesson extract III

Presenter We'll pick up the lesson again in its final section, when the pupils were being asked to work out what the English ought to have done.

INSERT 5D

Lesson extract IV

Presenter You can see how this problem-solving approach enables the teacher to pose a great many questions demanding analysis and reasoning.

And you will have noticed that the length and complexity of responses increased whenever the teacher asked this type of question. It's important to note that you can vary the level of your questioning, whatever the ability of your class. Knowledge questions can be made easy or difficult, to suit the calibre of your pupils; and even children of low ability can be stimulated to reason and make judgements, at a simple level.

To sum up -

(1) Avoid the over-use of one-word answer questions

- they restrict thinking
- they give no verbal practice

(2)/

(2) Make a deliberate effort to increase the proportion of questions demanding analysis and reasoning.

In this way, you should involve your pupils at a higher level of mental activity. Your own role will become less dominant, and the pupils will take a more responsible share in their own learning.

FADE

APPENDIX H

Model Script
Dealing with Answers

Rehearse and Record

15 November '72 (teaching sequences)
1330 - 1630

11 December '72 (presenter's links)
1330 - 1630

Teacher's voice

example of probing

(Beginning Insert 3)

FADE

Presenter Whenever you ask questions demanding thought, you must give pupils time to work out an adequate response. The teacher who is in such a hurry to get on that he abandons his own questions or answers them himself, without giving anyone the time to think out a good reply, soon discourages his pupils from making much mental effort.

"Why bother," they might say, "if he's not really interested in listening to our replies?"

You should practise pausing, after questions that demand an effort of thought. But let the class know that you expect them to use the pause, not just to relax through it. Note how an experienced teacher uses cues to make his intention clear to the pupils. The lesson is about the way that the Romans spread their power over the surrounding tribes.

INSERT/

INSERT 1

Teacher demonstrates pausing and cueing.

Presenter If a pupil's first attempt at an answer is inaccurate, incomplete, muddled or incoherent, your first reaction may be to reject it out of hand; or you may be tempted to accept the answer and perhaps improve on it yourself. Neither tactic helps the pupil's learning. You can often help the pupil to do better by the technique of prompting.

You can

- re-phrase your own question in a simpler form;
- provide hints that lead the pupil to greater accuracy;
- encourage the pupil to repeat his response more coherently.

Through prompting, a teacher can help pupils to achieve a higher standard of answering.

INSERT 2

Teacher using prompting.

Presenter Even when an initial response is acceptable, you may wish the pupil to extend, clarify, or justify his answer. You can encourage him to do this by probing his response, using supplementary questions. Through probing, you can control the difficulty of the pupils' learning tasks. Note the ways in which this teacher uses probing to raise the level of response.

INSERT 3

Teacher using probing.

Presenter When you ask questions demanding thought and reasoning, you may find yourself concentrating on a few mentally active pupils, while the rest of the class switches off. You should always try to involve as many pupils as possible in your questioning; and one way of achieving this is to re-direct initial answers to other pupils, asking them to comment. You can use phrases such as

- Would you agree?
- What is your opinion of that?
- Who has a different point of view?

Re-direction of initial responses can help to build up a discussion. In this way, you can reinforce the social aspect of questioning, as the teacher is doing in this sequence.

INSERT 4

Teacher using re-direction.

Presenter The skills of pausing, prompting, probing and re-direction will help you to obtain adequate answers and to build on initial responses in a constructive manner. In this way, you can

- set high standards, involving a maximum effort of thinking on the pupils' part;
- encourage good verbal expression;
- develop/

- develop pupil-to-pupil interaction, thus exploiting the social element in questioning.

FADE

APPENDIX I

Jordanhill College History Dept.

Student Evaluation Form

Questioning Skills

Dealing with Answers

How to use this form

1. Read the form before you start viewing the micro-lesson.
2. While viewing, make brief notes relating to the skills being practised.
3. After viewing, complete the assessment columns, including the "overall assessment".

STUDENT-TEACHER DATE SUBJECT-MATTER TEACH/RE-TEACH	Weak E	Not very effective D	Reasonably effective C	Clearly effective B	Very good A
1. Pausing, to allow pupils time for thought in formulating answers.					
2. Use of prompting, to encourage adequately phrased responses.					
3. Use of probing, for clarification and extension of response.					
4. Use of re-direction, for comment and discussion.					

Notes on performance

OVERALL ASSESSMENT

E → A

APPENDIX J

Jordanhill College History Dept. Micro-teaching Research Project

Student Questionnaire

Term 2

(1) This questionnaire forms an important part of our attempt to evaluate the technique of micro-teaching. Please complete it as fully and objectively as you can.

(2) Completing the questionnaire should not take up too much of your time. Most of the questions can be answered with a tick (✓), though occasionally we invite you to make additional comments.

(3) We ask you to put your name on the questionnaire, to enable us to correlate your responses with other data in this research study. You can be assured, however,

a. that all information will be treated as strictly confidential

b. that all questionnaires will be destroyed when the research study is completed.

(4) It will help us if you complete the questionnaire as soon as possible after your micro-teaching session, while the experience is still fresh in your mind. Completed questionnaires should be returned to

Mr. D. MacIntyre
History Department (Room 622)
Jordanhill College

NAME DATE

SECTION A STUDENT HANDOUT

At the beginning of Term 1 you were given a handout called "Four Basic Questioning Skills". It contained an analysis of skills which we have been studying throughout the session and which you have been practising in your micro-teaching.

Please answer the following questions by ticking the response which you feel to be appropriate.

1. Did you find time to read the handout

a. before or during your Term 1 teaching practice?

- Yes	
- No	

b. before your Term 2 micro-teaching?

- Yes	
- No	

If you answered 'No' to both parts of question 1, turn to Section B. Otherwise, please answer questions A2 - 6.

---

2. Did you find the analysis of questioning skills in the handout

- obscure in itself and of very little help to you in your teaching?	
- clearly set out and of practical use to you in your teaching?	
- clearly set out, but too theoretical to be of practical help to you in your teaching?	
- inadequately explained, but apparently relevant to your teaching?	

3. After your initial reading of the handout, did you refer back to it during your school teaching practice

- occasionally?	
- frequently?	
- not at all?	

4. In particular, did you refer to the handout when you were planning the lessons that you taught for assessment by your tutor?

- Yes	
- No	

5. If you disagreed with any statements in the handout, please give details.

6. If you would like to make further comments on the handout, please do so below

SECTION B THE MODEL VIDEOTAPES

(Please tick the responses which you feel to be appropriate)
At the start of each day's micro-teaching, you viewed and briefly discussed a videotape showing an experienced teacher demonstrating the use of the skills that you were about to practise.

1. Did you feel that the videotape models

- made the nature of the skills clearer to you than they had been before?	
- added nothing to your understanding of the skills?	
- left you more confused about the nature of the skills than you had been before?	

2. The videotape models consisted of brief sequences of teaching linked by explanatory commentary. Did you feel that the commentary was

- essential to your understanding of the skills being demonstrated?	
- acceptable as a reinforcement of the points demonstrated in the teaching sequences?	
- taking up time in the tape which could have been better spent in developing the teaching sequences at greater length?	

3. The time allowed for viewing and discussing each videotape model was half an hour. Relating this to the time available for the micro-teaching sessions as a whole, did you feel

- that the time spent on the models was about right?	
- that you could profitably have spent longer over viewing and discussing the models?	

4. The viewing of the models took place immediately before your micro-teaching practice. Did you feel

- that this immediacy was to be preferred, since it enabled you to practise the skills while the model was still fresh in your mind?	
- that you would have preferred a gap of a day or two between viewing the model and doing your micro-teaching, so that you could incorporate the advice of the model in your own lesson planning?	

5. The complete micro-teaching session consisted of

- planning the lesson
- viewing the model
- teaching
- review and re-plan
- re-teaching
- review

In the context of the micro-teaching session as a whole, did you find the videotape model component

- most useful of all?	
- very useful?	
- reasonably useful?	
- not very useful?	
- no use at all?	

6. If you have further comments that you would like to make about the videotape models, please add them below.

SECTION C THE MICRO-TEACHING PRACTICE

Please tick the responses which you feel to be appropriate.

1. Planning

a. Did you find the briefing letter sent to you

- all that you needed as a basis for planning?	
- reasonably adequate, but more details would have been appreciated?	
- inadequate as a basis for planning?	

b. How much time did you spend in planning your initial micro-teaching lessons?

Lesson A (Asking Questions)	- less than $\frac{1}{2}$ hour	
	- $\frac{1}{2}$ hour - 1 hour	
	- 1 hour - 2 hours	
	- more than 2 hours	

Lesson B (Dealing with Answers)

- less than $\frac{1}{2}$ hour	
- $\frac{1}{2}$ hour - 1 hour	
- 1 hour - 2 hours	
- more than 2 hours	

c. Did you find the planning of your micro-teaching lessons

- relatively easy?	
- rather difficult?	

d. If you ticked "rather difficult", was the difficulty mainly due to

- lack of guidance in choice of subject-matter?	
- problems of presenting and developing the subject-matter in 7-8 minutes?	
- having to concentrate in your lesson plan on a few specific skills?	

(n.b. You may tick more than one answer if you wish.)

2. Teaching

a. The following comments typify a range of reactions to the experience of micro-teaching. From the experience of teaching your own micro-lessons, tick those comments with which you agree. (n.b. You may tick several comments if you wish.)

- The handful of pupils made the whole experience seem artificial to me.	
- Having only a few pupils enabled me to concentrate upon particular teaching skills, without having to worry about problems of discipline or organisation.	
- The 7 - 8 minutes of teaching was long enough for me to cover the particular skills that I was practising.	
- The teaching period was too short for me to present the subject-matter properly or practise the skills effectively.	
- I found that concentrating on specific skills robbed me of spontaneity as a teacher and made me feel self-conscious.	

2a. (contd.)

- Concentrating on a few skills at a time forced me to think hard about them and helped me to make a deliberate attempt to practise them.	
- The micro-teaching was quite pleasant, but I can't see the relevance to normal classroom teaching.	
- I think the micro-lessons are a useful preparation for practising the same skills in the classroom.	
- I was distracted by the microphones and tape-recorder.	
- I didn't have a clear idea of the skills I was supposed to be practising.	

b. If your dominant reactions to micro-teaching are not covered by the statements in a. above, please add your own comments.

2. c. In each of your micro-teaching lessons you were asked to concentrate on two fairly complex skills. Did you feel

- that you were able to cope with both skills adequately in the one lesson?	
- that you would have preferred to concentrate on only one aspect of a skill (e.g. <u>either</u> prompting <u>or</u> probing <u>or</u> re-direction) in any one lesson?	

3. <u>Review and Re-plan</u>

a. The review of your micro-teaching was run as a group session. Please indicate your reactions by ticking the comments with which you agree. (n.b. You may tick more than one if you wish.)

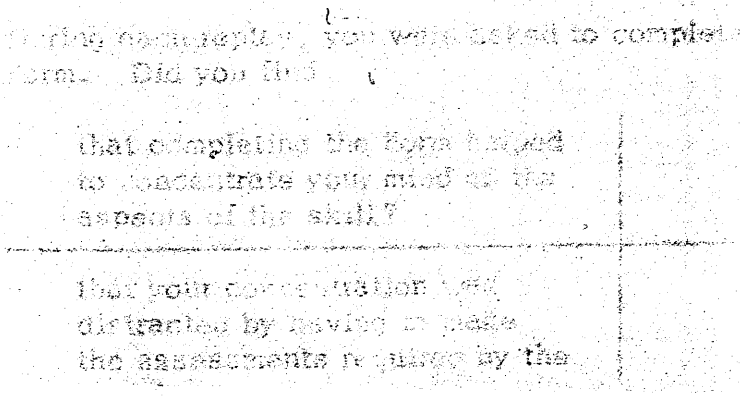
- I found it helpful to have comments on my teaching from other students as well as the tutor.	
- Listening to other students' teaching as well as my own was a useful experience.	
- I found the presence of the other students embarrassing when my tape was being replayed, and I should have preferred to review the tape on my own or with the tutor alone.	
- No strong feelings. The presence of other students was neither embarrassing nor stimulating.	

Please add any other comments if you wish.

3. b. A tutor was present during your micro-teaching and at each review session. Please indicate your reactions by ticking the comments with which you agree. (n.b. You may tick more than one comment if you wish.)

- His presence at both teaching and review was essential to the value of the exercise.	
- I should have preferred to do the teaching on my own, but I welcomed his guidance at the review sessions.	
- I was embarrassed by the presence of the tutor, and I think I should have been able to concentrate better on my teaching and evaluation if he had not been there.	
- The tutor's presence did not inhibit me, but I think the exercise would have been just as valuable if he had not been there.	

Please add any other comments if you wish.



3. c. Allowing for the brief time available for discussion, how would you rate the rôle that the tutor played in your review sessions ?

- I felt the tutor was too authoritarian in his criticism. He should have let the tape recording speak for itself and left me to make my own judgment.	
- I should have welcomed more positive guidance and criticism than I received.	
- His comments were very helpful in drawing my attention to aspects of the skill on which I needed to concentrate.	

Please add any other comments you wish to make on the tutor's rôle.

d. During each replay, you were asked to complete an Evaluation Form. Did you find

- that completing the Form helped to concentrate your mind on the aspects of the skill ?	
- that your concentration was distracted by having to make the assessments required by the Form ?	

3. e. Immediately after the replay of your tape, you were asked to comment on your own performance. Did you

- find this an embarrassing ordeal?	
- experience difficulty in thinking and speaking objectively about your own performance?	
- welcome the chance of criticising yourself instead of being criticised?	
- feel that the time would have been better spent, from your point of view, in listening to the comments of the tutor and the other students?	

(n.b. You may tick more than one comment if you wish).

f. In general, did you feel that the review sessions were

- valuable and about the right length?	
- valuable, but they should have been longer?	
- valuable, but they could have been shortened?	
- not very valuable?	

4. Re-teaching

a. Between teach and re-teach, you were given the opportunity to replan your lesson. Did you in fact

- make major changes in your lesson plan and/or the form of your questions?	
- make a few deliberate modifications to correct weaknesses apparent during the review?	
- decide to stick to the same plan, but to adapt in the course of the re-teach to the differing reactions of a new set of pupils?	

4. (contd.)

b. Did you feel that the re-teach lesson

- helped to improve your confidence in using the skills?	
- prolonged the micro-teaching session without adding anything of value?	
- enabled you to correct weaknesses which you had noted in your initial teaching?	
- had little effect on your ability to use the skill you were practising?	

(n.b. You may tick more than one comment if you wish.)

SECTION D GENERAL COMMENTS

Please add any further comments you would like to make, either to elaborate on a previous response or to make additional points.

APPENDIX K

Jordanhill College History Dept. Micro-teaching Research Project

Final Student Questionnaire

1. This questionnaire is a follow-up to the one you completed immediately after your micro-teaching sessions in Term 2. We should now be grateful to have your final appraisal of micro-teaching, in the light of your training experience in general and your Term 3 experience of teaching practice in particular.

2. As before, the questions can be answered with a tick (✓), but we should also welcome any additional comments you may care to make in the space provided.

3. Please complete this questionnaire as fully and objectively as you can, add your name (so that we can relate this appraisal to your previous comments) and return the completed form to

> Mr. D. MacIntyre,
> History Department (Room 622)
> Jordanhill College.

4. Once again, we should like to assure you that the questionnaires will be treated as strictly confidential and that they will be destroyed immediately the research study is completed.

NAME DATE

SECTION A — APPLICATION OF SKILLS TO CLASSROOM PRACTICE

1. The seven questioning skills which you practised in micro-teaching sessions were

   a. Avoiding over-use of "one-word-answer" questions

   b. Avoiding over-use of "yes/no" questions

   c. Varying levels of questions, with optimum use of questions calling for analysis, reasoning, judgment, etc.

   d. Encouraging adequate response through pausing and cueing, to give pupils time to think out well formulated answers

   e. Encouraging adequate response through prompting, where a pupil either fails to reply or makes an inadequate, inaccurate or incoherent response

   f. Use of probing, to extend, clarify or justify initial response

   g. Re-direction of responses to other pupils, for comment and discussion.

To what extent did you make a conscious attempt to apply these skills in your teaching practice subsequent to the micro-teaching sessions? (Please tick whichever response you feel to be most appropriate.)

Not at all	
Occasionally, when I remembered	
Consistently, whenever I employed question techniques	

2. Please indicate, by ticks in the appropriate columns, the degree of difficulty or success which you experienced in attempting to apply the seven skills.

	Very difficult to apply	Rather difficult to apply	Moderate success	Good success
Restricting "one-word-answer" questions				
Restricting "yes/no" questions				
Optimum use of "higher order" questions				
Pausing				
Prompting				
Probing				
Re-direction				

3. If you found the skills difficult to apply, was it mainly because

- use of the skills conflicted with your normal teaching style and personality ?	
- these question techniques failed to produce a satisfactory level of response from the pupils ?	
- other factors, e.g. disciplinary problems, size of class, etc., inhibited the use of the skills ?	

Please comment on any other problems you encountered in attempting to apply these skills

4. What effect, if any, did your application of these skills have on the length and quality of the pupils' responses?

- it made no noticeable difference	
- I detected a slight improvement	
- there was a clearly marked improvement	

5. How would you assess the problem of transferring the use of these skills from the small group teaching situation (i.e. micro-teaching) to the full-scale class lesson?

- I found it impossible to make the transfer	
- it was difficult, but the transfer could be achieved with practice	
- I found no difficulty in making the transfer	

SECTION B GENERAL COMMENTS ON MICRO-TEACHING

1. In retrospect, how would you assess the value of micro-teaching as an element in your training course, along with teaching practice, lectures, seminars and tutorials? (Please tick the response which you feel to be most appropriate.)

- of great value. More time could be spent on it with advantage	
- interesting, but it did not help me to cope any better with the practical problems of classroom teaching	
- of reasonable value. About the right amount of time was spent on it, in relation to the other elements of the course	
- of no value at all	

2. Would you recommend the retention of some form of micro-teaching as an element in the course work of graduates in future sessions?

Yes	
No	

3. If you answered "yes" to question 2, do you think that micro-teaching should be applied

- as a form of initial training for all graduates, before they go out on teaching practice?	
- as a reinforcement of practical experience offered to all graduates at the mid-point of the training year?	
- as a form of remedial training for those graduates showing weakness in specific teaching skills?	

(N.B. (i) You may tick more than one response if you wish.

(ii) Remember that micro-teaching can apply to a variety of teaching skills, not simply to question techniques.)

4. There are three main elements of micro-teaching:

a. the definition and analysis of teaching skills

b. the practice of those skills under controlled conditions

c. self-assessment based on audio or audio-visual feedback

Clearly, one could include in a training course an element of a. without b. or c.; or an element of a. and b. without c. Given the constraints of time in the one-year graduate training course and considering the balance and integration of the course as a whole, would you wish to include

- an element of a., without b. or c.?	
- elements of a. and b., but not c.?	
- elements of a., b. and c., as a complete micro-teaching "cycle"?	

5. Please add any other comments you may wish to make about micro-teaching, with particular reference to its possible applications and modifications in future graduate teacher training courses.

APPENDIX L

Jordanhill College History Dept. Micro-teaching Research Project

Lecturers' Questionnaire

1. This questionnaire seeks evidence of your reactions to the concept and practice of micro-teaching as it has involved you during 1971-73. It also seeks your recommendations for possible future developments in this technique of training.

2. The evidence from these questionnaires will form an important complement to our measurements of student attainments during 1972-73. I shall be most grateful if you can find time to complete the questions as fully and frankly as possible.

3. Adding your name to your completed questionnaire will provide guidance for future discussions within the History Department. You can be assured, however, that no names will be attached to any summary of opinion included in public reports or accounts of this micro-teaching project.

4. Please return your completed questionnaires to me.

NAME

D.C. Butts
Co-ordinator,
Micro-teaching Project.

June 1973

SECTION A — THE CONCEPTUAL BASIS OF MICRO-TEACHING

1. The technique of micro-teaching is based on the belief that the "art of teaching" can be broken down into identifiable skills and behaviours, each of them capable of close analysis. It is further assumed that the student can improve his teaching skills by

   a. studying this analysis in some detail

   b. practising selected skills under controlled conditions that allow him to concentrate for short periods on one or two skills at a time

   c. assessing his own performance.

The following comments represent a range of reactions to the beliefs set out above. Please tick those comments with which you broadly agree.

- Effective teaching, assessed in terms of pupil learning, is associated with many different teaching styles which cannot be defined in terms of specific skills.	
- Specific teaching skills and behaviours can be identified, analysed and consciously applied.	
- Teaching can be subjected to a theoretical analysis of skills and behaviours, but deliberate practice of these skills under controlled conditions is unlikely to lead to increased effectiveness in the classroom.	
- Specific teaching skills can be strengthened through deliberate practice under controlled conditions.	
- In developing his skills, the teacher must rely on personality and intuition rather than self-conscious techniques.	

If your reactions are not covered by these statements, please add your own comments.

2. In this project, fourteen teaching skills were distinguished and grouped under the four broad headings of Preparation, Presentation, Pupil Involvement and Relationships (see "Analysis of Teaching Skills"). If you were making your own analysis of teaching skills and behaviours,

- which if any of the fourteen skills would you omit?

- which if any skills would you add to the list?

SECTION B USING THE APPRAISAL GUIDE

Presumably, in making your assessments of students visited on teaching practice, you take account of

a. the quality and effectiveness of the teaching in general

b. particular strengths and weaknesses displayed at various points of the lesson.

1. Did you feel on your teaching practice visits this session that the use of the Appraisal Guide (together with the Analysis of Skills on which it was based) helped to make your assessments

	Yes	No
- more systematic		
- more precise and specific		
- more objective		
- more reliable (in the sense of applying the same criteria to all students visited)?		

2. Did you feel that using the Appraisal Guide (including the necessity of completing it)

	Yes	No
- hindered you in your task of making a balanced assessment of the student?		
- tended to channel your observation in predetermined directions, making it difficult to pay adequate attention to details of behaviour not specifically mentioned in the Guide?		

3. In addition to providing for an assessment of general teaching skills, the Appraisal Guide asked for a more specific assessment of seven aspects of questioning technique. Did you find that the task of making both general and specific assessments within the same lesson was

- fairly simple	
- impossible to carry out adequately	
- difficult, but not impossible?	

4. An alternative method of assessing a specific skill within a lesson would be for the observer to set aside a short period (say ten minutes) of the lesson and to concentrate for that period exclusively on measuring the skill, perhaps employing some form of quantitative analysis. The remainder of the lesson would then be devoted to general assessment. Please comment on this suggestion, in comparison to the method employed this session. (Better? Worse? More or less practicable? etc.)

5. During the three group assessments of video-taped student lessons carried out in 1971-72, it was found difficult to achieve a high standard of reliability among lecturers' assessments. Discrepancies were most marked when it came to detailed assessments of specific skills. Which of the following factors do you feel contributed to this lack of reliability?

- Difficulty (in spite of the Skills Analysis) in arriving at an agreed concept of each skill	
- Difficulty in noting and assessing a variety of different skills within the one lesson	
- Difficulty in agreeing upon a concept of "average" performance	
- Difficulty in agreeing upon a distribution pattern, ranging through seven grades from "Weak" to "Very good"	
- Difficulty in making reliable allowance for "appropriateness" and "grasping of opportunities" in using the skills (cf. my letter of 2-10-72)	

Please mention any other contributing factors not included in the above list.

6. What is your reaction to the two following criticisms of the Appraisal Guide in its present form?

a. As an instrument for measuring teacher effectiveness, the Appraisal Guide leads the observer into a no-man's land between scientific analysis and intuitive assessment.

Agree	
Disagree	

Comment if you wish.

b. The Appraisal Guide obliges the observer to take a teacher-centred view of the lesson situation. An instrument which took more account of pupil reaction and achievement would provide for a more adequate assessment of teacher effectiveness.

Agree	
Disagree	

Comment if you wish.

7. Quite apart from any decision which your department may take to continue experiments in micro-teaching, would you wish to continue regular use of some structured form of assessment based on a detailed analysis of skills and behaviours ?

Yes	
No	

SECTION C — FUTURE DEVELOPMENTS IN MICRO-TEACHING

1. From your experience of micro-teaching so far, do you feel that it should be persisted with in any form in the training of history graduates?

Yes	
No	

2. If your answer to question 1. was "yes", would you wish to see an element of micro-teaching incorporated

a. at an initial stage of the course (for all students)

(i)	to give basic training in specified skills*	
(ii)	to provide a basis for general comment by tutor and student*	

(*The brief for (i) would be "Practise skill x in a 10-minute lesson". The brief for (ii) would be "Teach a 10-minute lesson on a given topic".)

b. at a later stage of the course (for selected students)

(i)	to give supportive training to students showing weakness in previously identified specific skills	
(ii)	to give supportive training to students classified as "generally weak"	

3. Assuming that your answer to question 1. was "yes",

   a. How much time (number of days) should be given to micro-teaching for any one student?

   b. Should this time be taken from teaching practice or from in-college courses?

4. Bearing in mind the practical problems of time, transport, accommodation, etc., would you prefer micro-teaching to be

- carried out in schools (perhaps with a number of students gathering in one school)?	
- carried out in college (involving transport of pupils to and from college)?	

What would appear to you to be the major practical problems involved in introducing micro-teaching on a regular basis?

5. Do you think that micro-teaching (if introduced on a regular basis) should be

- a shared responsibility among all lecturers?	
- a responsibility allocated to a small team of lecturers within the department?	

Please comment if you wish.

6. Assuming the introduction of some form of three-phase training along the lines of the G.T.C. recommendations for the training of secondary teachers, in which phase(s) should an element of micro-teaching be incorporated?

Phase I	
Phase II	
Phase III	

Please comment if you wish.

APPENDIX M(a)

Hints for completing the Appraisal Guide (letter to tutors, October, 1972)

Here is a summary of the points on which we have agreed during discussions.

(a) On each visit, establish a concept of the weak → average → very good lesson. This concept will obviously be shaped by consideration of - the age and ability of the class
- the content chosen for the lesson
- the teacher's objectives.

(b) Use this weak → average → very good continuum as a yardstick for measuring the student's performance. This, of course, is what you would be doing in any case, in determining your "official" assessment. However, in using the Appraisal Guide, you should not make allowances for the student. Your yardstick should be firmly and objectively based on your concept of what the good lesson should be like, and considerations of "what kind of lesson could any student be expected to teach at this stage?" or "what personal problems did the student have to overcome?" should not affect your judgment. The objective "good lesson" criterion should be applied to each assessment that you make on the Appraisal Guide.

(c) This objectivity does not mean that you make a mechanistic measurement of the display of any particular skill or behaviour, regardless of the circumstances of the lesson. If the skills practised in micro-teaching sessions are to have any relevance to effective teaching, they must be capable of "selective transfer" to the classroom situation. It is for this reason that we have included words like "over-use" and "optimum"/

"optimum" in our analysis of the skills. It is therefore essential that you should assess each student on the basis of his appropriate use of the skills - appropriate to the age and ability of the class, the objectives and content of the lesson, and so on.

(d) Having said that, it is assumed that you agree that all the specific skills selected for study are appropriate in some measure to the great majority of class-teaching lessons. Students have been given an analysis of these skills, and they have been told that lecturers will be paying particular attention to them when observing lessons.

(e) Scoring the seven aspects of the "specific skills for micro-teaching" demands a more analytical approach than you might normally make to an assessment of question technique. It goes without saying, however, that, since your main pre-occupation will be with helping the student, you will not wish to concentrate on the specific micro-teaching skills at the expense of your judgment of the lesson as a whole. Provided that you have the seven points of the questioning skills firmly in your minds, it should be possible to combine a balanced over-view of the lesson with a reasonably reliable assessment of the specific behaviours which we are studying.

(f) If you feel, after observing a lesson, that a student's omission of a particular questioning skill was due, not to a failure to grasp the opportunities which a good teacher would have taken, but to the fact that the use of the skill was simply not appropriate to the lesson, do not attempt to score that particular aspect. It will be a useful guide to me if you mark the row of boxes thus:

E	D	C-	C	C+	B	A
			N/A			

However, for the reasons stated in (d) above, this situation should not occur very often.

(g) At the risk of re-stating the obvious, the following notes are set out as a guide to establishing norms for assessment.

1. Avoiding over-use of one-word answer questions

Below average - Clearly uses a disproportionate number of these questions. Does not appear interested in obtaining adequate, "complete sentence" answers from pupils.

Average - Uses rather too many of this type of question. Occasionally phrases questions in a form designed to produce a more complete response, but you feel he could do more in this direction.

Above average - Makes legitimate use of one-word answer questions but, wherever possible, phrases his questions in a form designed to produce a more complete response. You have the feeling that he is deliberately trying to encourage adequately phrased responses from the pupils.

2. Avoiding over-use of yes/no type questions.

As for 1. above.

N.B. The student who follows up his yes/no questions by an "analysis probe"/

probe" seeking justification of the initial response is clearly teaching more effectively than the student who fails to do this; but repeated use of this formula may still be thought to constitute over-use of the yes/no type question.

e.g. T. Do you think that James IV was an able king?

P. Yes.

T. Why? What reasons can you give?

This kind of exchange could be avoided by phrasing the initial question in an analysis form.

3. Optimum use of questions demanding interpretation, judgment, etc.

Below average - Takes very few of his opportunities for asking questions of this type.

Average - Takes some of his opportunities but you feel he could do more.

Above average - Clearly takes deliberate steps to obtain thoughtful responses at all appropriate points of the lesson.

4. Giving pupils time to think out adequate answers

Below average - A perfunctory approach to questioning. So keen to press on that he either answers his own questions or abandons them, before pupils have reasonable time to think out a reply.

Average - Occasionally gives time for thought, but is not consistent about this.

Above/

Above average - Makes us feel that, when he asks a question demanding thought, he is really interested in the quality of the response. He therefore paces his questioning appropriately, and, through cueing, ("Now, think carefully about this", etc.)

encourages the pupils to give good replies.

5. Use of prompting, to encourage well-formulated answers

Below average - accepts or rejects inadequate answers without comment. Does not try to help pupils to improve upon inaccurate or incomplete initial responses.

Average - Remembers occasionally to prompt, but does not take all his opportunities.

Above average - Shows patience and persistence in prompting, making it clear that he is concerned both to help the pupils and to set a high standard of response.

6. Constructive use of answers through probing (for clarification, extension of response)

Below average - Rarely takes his opportunities to extend the pupils' thinking through this type of follow-up question.

Average - Uses this technique occasionally, but could appropriately do more.

Above/

Above average - Takes most of his opportunities to obtain maximum effort of thinking from pupils by use of this technique. His probing questions are in a form well calculated to stimulate further response.

7. Development of discussion, through re-direction of initial answers for further comment

Training in this skill assumes the value of extending pupil involvement by establishing pupil - to - pupil communication as well as pupil - to - teacher responses. It aims at producing a situation in which pupils feel that they have some share in getting at the truth; and that the teacher does not necessarily have the monopoly of the right answers.

Obviously, any teacher with a scheme of work to follow has to use this technique with discretion. Its most effective use is perhaps within the context of open-ended, "thought" questioning.

Below average - Rarely, if ever, uses this technique.

Average - Uses the technique occasionally, but does not exploit it to full effect.

Above average - Makes the most of his appropriate opportunities. His general relationship with the pupils gives them confidence in voicing their own views, arguments, judgments, etc.

APPENDIX M(b)

Supplementary note on avoidance behaviours (December, 1972)

One or two lecturers have drawn attention to the difficulty of making a realistic assessment of "avoiding over-use of single-word answer and yes/no questions" in lessons where the student asks very few questions of any kind. In these circumstances, it is clearly illogical and misleading to give the student a high assessment on these two "avoidance" behaviours. The best procedure is to mark the behaviours as "not applicable", adding a brief note of explanation (e.g. "student asked very few questions"). The student's ability on questioning will then be accurately reflected in your assessments of the other five skills. (These will probably be low assessments, since in most cases the student will have failed to use the opportunities for questioning which the lesson topic provided).

APPENDIX M(c)

Reminder letter
(February, 1973)

It would be helpful if, before commencing your Term 3 visits, you could refer again to paragraph 5 of my letter dated 2 October 1972 and to my follow-up letter dated 14 December 1972. In particular, I should like to stress the importance of

(i) establishing an objective weak/average/good lesson continuum in your mind as a yardstick for assessing performance;

(ii) basing your judgments on the extent to which the student took his opportunities to make appropriate use of each skill;

(iii) making a "not appropriate" return in those instances where you feel that a skill remained unpractised, not because the student failed to grasp his opportunities, but because the nature of the lesson and/or the characteristics of the pupils did not allow for its practice;

(iv) making a "not appropriate" return for the two "avoidance behaviours" (single-word answer and yes/no questions), if you feel that "avoidance" arose simply from failure to ask many questions of any kind.

APPENDIX N(a)

Term 1 briefing letter to students

Subject: This can be of your own choosing, preferably to fit in with the scheme of work which the class is following. The lesson should, however, lend itself naturally to a period of questioning and discussion. You can check your choice of subject with your Principal History teacher.

Duration: The lesson should be planned to last 25 minutes. This should give me plenty of time to set up and dismantle my recording equipment at the beginning and end of the period.

Structure: Plan your lesson to include

- approximately 10 minutes of introduction and exposition
- approximately 10 minutes in which you question the class and do your best to develop a discussion based on and maintained by your questions
- approximately 5 minutes of summing up.

Lesson Plan: You should prepare a lesson-plan and let me have a copy when I come to record your lesson. The plan need not be elaborate, but it should indicate clearly the three main divisions of your lesson, so that I can judge the points at which you are moving from one stage to the next.

Recording Equipment: This will be kept as simple as possible, and should in no way inhibit you or your pupils. You can plan and teach your lesson in the normal way.

Evaluation:/

Evaluation: I should like to stress once again that any analysis that I make of your recorded teaching will be purely for research purposes and will in no way be taken into account in your "official" teaching assessment. All recordings will be wiped at the end of the session.

APPENDIX N(b)

Term 3 briefing letter to students

Class: I have arranged that you should teach a first-year class as you did before.

Structure of lesson: This should follow, as closely as possible, the structure of the lesson which you recorded in Term 1. That is to say, the lesson should

- last approximately 25 minutes;
- be based on a subject and approach which lends itself naturally to a period of intensive questioning and discussion;
- be planned in three main stages;
  - (i) approximately 10 minutes of introduction and exposition
  - (ii) not less than 10 minutes in which you question the class and do your best to develop a discussion based on and maintained by your questions
  - (iii) approximately 5 minutes of summing up.

Lesson notes: As before, I should like you to prepare brief lesson notes which indicate clearly the points at which you move from one stage to the next. (This enables me to identify the "questioning" period for the purpose of analysis).

Skills covered: I shall be paying particular attention to the skills described in your hand-out "Four Basic Questioning Skills", i.e.

1. Avoiding over-use of "one-word-answer", "yes/no" questions;
2. Making optimum use of "higher order" questions;
3. Encouraging adequate answers through cueing and prompting;
4. Making constructive use of initial answers through probing and re-direction.

Finally/

Finally, a reminder and a reassurance:

(a) Please do not discuss your participation in the micro-teaching project with your history tutors, unless problems over visiting dates force you to do so. In particular, it is most important that students who received micro-teaching practice in Term 2 should not mention the fact to their tutors.

(b) My assessments are for research purposes only. No one but myself will listen to the audio tapes and they will be wiped at the conclusion of the study.

Jordanhill College History Dept. Micro-teaching Project

Ground Rules for Analysis and Coding of Student Lesson Excerpts

1. General Aims and Procedure

(a) The analysis and coding procedures are designed to provide a quantitative measurement of students' use of the seven questioning skills selected for the micro-teaching exercises. These skills are -

- avoiding over-use of one-word-answer type questions
- avoiding over-use of yes/no type questions
- making optimum use of higher order questions
- giving pupils time to think out adequate answers
- prompting, to encourage accurate, well-formulated answers
- probing, to extend, clarify and justify initial responses
- re-direction of initial responses to other pupils, for comment and discussion

(b) The analysis and coding will be applied to selected ten-minute excerpts of classroom lessons taught by students in terms 1 and 3; and to selected excerpts of micro-teaching lessons taught by students in term 2.

(c)/

(c) Each lesson sequence will be subjected to two separate analyses:

(i) to determine the percentage of the lesson period devoted to

- pausing to allow pupils time to respond to questions
- pupils' verbal responses to questions

(ii) to determine the types of questions asked by the teacher and to note failure or lack of opportunity to respond. Coding will cover -

- higher, middle and lower order questions
- single-word answer and yes/no answer questions
- questions eliciting no pupil response
- questions where no opportunity is given to respond
- prompts, probes and re-directions

2. Measuring the time devoted to pupil response

(a) Since pupils' responses are often very brief and since the teacher's voice dominates most sound recordings, it has been found that the most accurate method of measuring pupils' responses and time allowed for response is to measure the extent of all other talk and activity and to subtract this figure from the total length of the excerpt under examination.

(b) Select the lesson excerpt to be analysed and time it exactly. Make a note of IN and OUT cues for reference.

(c) If possible, use the same tape recorder and the same stop-watch for all excerpts throughout the entire experiment.

(d) For each excerpt, make a cumulative timing of

- the length of teacher talk
- the length of teacher activity (e.g. writing on the blackboard, operating a projector) and of teacher-directed, non-verbal pupil activity (e.g. taking out jotters, turning to a page in a text-book)

(e) Distinguish between

(i) 'Natural' pauses by the teacher in mid-speech. These are to be counted as teacher talk - DO NOT STOP THE WATCH. (n.b. These pauses do not normally exceed two seconds.)

(ii)/

(ii) Pauses following a teacher question, where the teacher is waiting for pupils to formulate a response. STOP THE WATCH during these pauses, even if no response is forthcoming and the teacher subsequently resumes talking.

e.g. T. Can anyone tell me where the Romans first landed in Britain? ------ (STOP THE WATCH) ------ Come on, you should know the answer to this one ------ (STOP THE WATCH) ------ Well, I'll give you a clue, etc.

(f) When the voice of the teacher and pupils overlap, measure the duration of all teacher talk.

3. Definition of a teacher question

A teacher question is defined as a word, a group of words or a unit of non-verbal behaviour, having the apparent intention of eliciting a response from the pupils.

Teacher questions can be

(a) interrogative

e.g. Who won the battle of Stirling Bridge?

(b) imperative

e.g. Give me the name of the English commander at Stirling Bridge.

(c) implicit

Implicit questions can take several forms, e.g.

(i) a pause (see 5(e), usually following an initial response, in which the teacher says nothing, but indicates, through facial expression, gesture or even the mere act of pausing, that he is expecting further responses.

(ii) a comment on an initial response, conveying to the pupils the feeling that the original question is still "hanging" and that the teacher is expecting further responses. This expectation is indicated by a pause after the comment or by an inflexion of voice.

e.g./

e.g. a. T. What steps did Agricola take to stop the invasions of the northern tribes?

P1. He sent legions into Scotland.

T. Well, that's part of the answer. (Pause)

P2. He built a wall and a line of forts.

T. Yes, that's right.

b. T. What did Skara Brae people do to help keep the wind out of their houses?

P1. They lit fires inside the houses.

T. I don't think that would have helped much. (Pause)

P2. They filled the gaps in the walls with small stones.

T. Yes.

n.b. In each of the examples in (ii) above, the teacher comment (taken together with the pause) should be regarded and coded as a second asking of the question. The coder should distinguish between those instances where a teacher appears to be soliciting a sequence of responses to an initial question and those where a teacher is merely acknowledging a request by a pupil to make an unsolicited comment. A series of solicitations should be recorded as a fresh asking of the initial question.

e.g./

3. (contd.)

e.g. T. What benefits did the Romans bring to Britain?

(several hands are raised)

Yes, Mary?

P1. They built roads.

T. That's right. (To P2) Yes?

P2. They built towns as well

T. Towns as well, yes. Jim?

P3. They brought peace.

T. Good. You can see that the Romans did a great deal of good. In fact - yes, Angus?

P4. I don't think the Britons benefited much.

In this example, the first three nominations should be recorded as questions. The final nomination, however, is merely the acceptance of an unsolicited comment, and no question should be recorded.

4. Question Types

A. Lower Order Questions

The following types of question should be regarded and coded as Lower Order questions:

(i) Questions requiring the pupil to provide factual information, based on a recall of knowledge previously acquired. Such questions may also involve straightforward recognition of material. Responding to these questions does not involve the pupil in any activity (e.g. description, translation, interpretation) going beyond recognition or recall.

e.g. - What was the date of the battle of Flodden?

- Who succeeded to the throne of England in 1603?

- Give me some examples of things which the Romans built in Britain.

- When did James I die?

- How do you spell 'monarchy'?

- What is this part of a castle called?
(Teacher points to detail in picture)

- What is the name of this town (river, range of mountains, etc.)
(Teacher points to detail on map)

n.b. In the last two examples above, it is assumed that the visual material on which the question is based is in a form which has been previously studied by the pupil and which is therefore immediately recognizable.

(ii)/

(ii) Questions requiring the pupil to give a straightforward definition of a word, without involving him in the explanation of a concept.

e.g. - What does the word 'cuirass' mean?

BUT NOT

- Can you define what is meant by 'feudal system'?

(iii) Questions seeking to establish the pupils' range of knowledge, understanding or experience (generally as a form of feedback to the teacher).

e.g. - Have you heard of the Chartists?

- Did you know that the Romans divided Gaul into three parts?

- Do you follow what I've just told you?

n.b. Questions of this type are normally couched in a form that requires a simple yes/no response.

(iv) Questions inviting the pupils to express opinions or make suggestions based simply on personal preference or immediate and unconsidered reaction. Such questions make no apparent demands for cognitive work on the pupils' part.

e.g. - Do you like the story of Alfred and the cakes?

- Do you think you would have liked Wallace if you had met him?

- Who prefers the Royalists to the Roundheads?

- What/

- What name shall we give to the main character in our history play?

- Do you think that was a true story?

(v) Questions inviting the pupils to make suggestions based simply on random guesswork.

e.g. - How long do you think a knight might stay in a church on the night of his vigil?

n.b. The coder must distinguish these 'random guesswork' questions from questions inviting suggestions or predictions based on information at the pupils' disposal. His decision must be based on the form of the question and a consideration of what has gone before in the lesson.)

B. Middle Order Questions

In general, Middle Order questions require the pupil to

- organize and make some use of the information at his disposal

- perceive the content and structure of material presented to him

- make suggestions based on the direct application of previously acquired knowledge to new situations.

Middle Order questions do not involve the pupil in an act of analysis as defined in 4(c).

Such/

Such questions may require the pupil to

(i) provide a description or account in his own words, based on material previously studied.

e.g. - Can you describe the layout of a typical Roman fort?

- Give me an account of what happened at the battle of Stirling Bridge.

(ii) "translate" (i.e. express the meaning of) material provided in symbolic form into a form which involves a verbal response.

e.g. - Can you describe what you see in this picture?

- What do these symbols on the magnetic board represent?

- Take a look at this graph and then tell me what happened to the population of Glasgow between 1800 and 1850.

(iii) make a straightforward comparison, based on information at the pupils' disposal.

e.g. - From the information given in this document, who worked longer hours, the mill-hands or the farm labourers?

n.b. Sometimes a teacher will seek a comparison based on information shich he <u>assumes</u> is at the pupils' disposal (i.e. he has not specifically given them the information in the lesson prior to the question.) Such a question should be coded as Middle Order if it is felt that the teacher is making such an assumption.

e.g. - Do you think that in the 12th century people were more religious than they are today?

This/

This type of question should be distinguished from questions involving a comparison based simply on personal preference.

e.g. - Do you think that long hair looks nicer than short hair? (see A(iv)).

(iv) provide an explanation, in his own words, of material being studied.

e.g. - Can you explain what the writer is getting at in his final paragraph?

- Can you define what is meant by the "feudal system"?

n.b. This is a Middle Order question if it involves the pupil in providing an explanation in his own words. If it merely requires him to repeat a definition given earlier in the lesson, the question becomes Lower Order.

(v) make a summary of material presented in verbal, written or graphic form, stating the main ideas or main features involved.

e.g. - What are the three main arguments that this writer puts forward to support his claim that young children should not be allowed to work in the mines?

- From what you've learnt in this lesson, tell me the main features of the enclosure movement in 18th century Scottish farming.

- From what I've told you in today's lesson, can you sum up the main benefits that the Romans brought to Britain?

- Read over the document that we've been studying and then give me a summary of it in your own words.

(vi)/

(vi) make predictions based directly upon a consideration of trends apparent in material being studied (i.e. extrapolation).

e.g. - Look at this graph of population growth in Glasgow between 1800 and 1830, and then tell me what you think the size of Glasgow's population would have been by 1840.

(vii) apply information, principles or rules already acquired, in order to make suggestions relating to new situations and specific problems.

e.g. - (in a lesson on Roman forts)
What would the Romans have used a watch tower for?

- (after discussing the location, age and construction of the Egyptian pyramids)
Would you have needed a large labour force to build the pyramids, do you think?

- (after a lesson on the first Reform Act)
Look at the rating details given in this document and tell me which of the householders would have had the vote after 1832.

- (after telling the pupils about the categories of offences in the 18th century punishable by transportation and describing an incident in modern times involving an offence of a similar nature)
What do you think would have happened to that man if he'd been living in the 18th century?

n.b. In each of the above examples, the application of knowledge, principles or rules should lead the pupil directly to a correct response, without involving him in any fresh act of analysis/synthesis as defined in 4C below.

C./

C. Higher Order Questions

In general, Higher Order questions require the pupil to carry out a process of analysis, possibly leading on to synthesis or evaluation. Analysis in this context is defined as the breaking down of material into its constituent parts and the perception of relationships and organization among those parts. Analysis involves the pupil in a process of reasoning, which can lead him to identify motives and causes, make inferences, draw conclusions and formulate judgements.

Synthesis is defined as the creation of a new pattern or structure from many elements on which the pupil has worked at a level of analysis and reasoning.

Evaluation is defined as the process of making reasoned judgements, based upon a critical examination and appraisal of relevant evidence.

Higher Order questions may require the pupil to

(i) identify motives or causes

e.g. - From what you've learnt about James IV, why do you think he chose to go to war against the English?

- What motives do you think Cromwell had for destroying so many churches?

n.b. If the pupils had previously been taught a list of motives, the question would be rated as Lower Order (recall of information).

The question involves analysis if the pupils have sufficient data about Cromwell to enable them to make a critical appraisal of his motives.

(...)

(ii) draw conclusions and make generalisations from evidence presented

e.g. - What can we conclude about the political sympathies of the man who wrote this document?

- From the evidence of Shaftesbury's Commissioners what can you say about the condition of children in the mines in 1840?

(iii) elucidate evidence to support generalisations

e.g. - What evidence do these old railway posters provide to support the suggestion that rail travel had become popular by the mid-19th century?

(iv) distinguish objective statements from subjective opinions in secondary source material

e.g. - From your knowledge of conditions in Britain after 1815, which parts of this document would you say represent a personal bias on the writer's part?

(v) identify elements of agreement or disagreement between two pieces of historical evidence or secondary source material

e.g. - From a reading of these two articles, do you think the writers were agreed about the educational needs of working-class children in 1870?

n.b. This type of question should be rated as Higher Order if it requires the pupil to "detect the relationship and organisation of the constituent parts" of the material. If it merely requires the

the pupil to detect agreement or disagreement between simple statements of facts or figures, the question should be rated as Middle Order.

(vi) Make inferences, develop implications, frame hypotheses or make predictions, based upon an analysis of given material and involving a logical development of consequences implicit in that material.

e.g. - What can we infer from these paintings about mediaeval attitudes to death?

- Given the pattern of events that we've established, what do you think the likely outcome would be?

- Assuming that the northern tribes had overrun the Roman walls, what do you think the effect on the Roman occupation of Britain would have been?

- If the Corn Laws had not been repealed, what would have been the result, do you think?

- Considering the plan of Napoleon's campaign that we've dealt with so far, what strategies might he have adopted for this coming battle?

n.b. These examples should be contrasted with those given in 4B(vii), and the coder should distinguish between suggestions which derive from a direct application of principles to a specific situation and predictions which involve a process of analysis and logical reasoning. Questions in the Higher Order category of prediction are always open-ended. Compare Coltham J B and Fines J ("Educational Objectives for the Study of History") p. 20:

"In/

> "In this cognitive behaviour, what is already apprehended is used as a taking-off point for some idea which is rooted but not present in the evidence being examined. 'Since I know this, then it might be that---' is the kind of thinking involved, and it represents a more positive and adventurous mental behaviour than that of comprehension, and with some imagining added. (This behaviour) in dealing with historical material must, however, be distinguished from fantasy."

"Fantasy", in the terms of these Ground Rules, would be categorized as Lower Order responses (see 4A(iv) and (v)).

(vii) solve problems demanding a solution original to the pupil and involving an analysis of the relevant factors

e.g. - How can we assess the reliability of this historical document?

- What could Agricola have done to deal with the warlike tribes from Scotland?

(viii) produce original communications, involving the assembly of elements from amny sources into a fresh pattern or structure based upon independent thought and creative speculation

e.g. - Supposing you were a child working in a 19th century cotton mill, what kind of account of your life would you give to an interested stranger?

(ix) give opinions about issues, based on reasoned judgement

e.g. - Do you think that Parliament made the right decision in condemning Charles I? Give me arguments to support your view.

(x)/

(x) judge the validity of ideas

e.g. - How would you assess the realism of the ideas on social planning put forward in Morris' "News from Nowhere"? Give me a reasoned judgement.

(xi) evaluate the merits of solutions to a problem

e.g. - Which of the plans put forward after 1815 for solving unemployment seems to you to have had most merit? I want you to justify your answer.

(xii) argue a conclusion

e.g. - What reasons would you give to support your contention that the Roman occupation did more harm than good in Britain?

D. Procedural and Rhetorical questions

(a) Procedural

A procedural question arises from organisational or managerial aspects of the teacher's role.

e.g. - How many of you have jotters?

- Would you fill up this space in the front row?

- Do you all understand what you have to do?

- Did you all hear John's reply?

- Would you repeat that?

n.b. A procedural question does not require the pupils to recall or do any work upon any information.

(b) Rhetorical

A rhetorical question is a question that appears to call for no response from the pupils. Often the inferred response is contained within the form of the question.

e.g. - Who would have thought Elizabeth would do a thing like that?

* - Cromwell would have taken that for granted, wouldn't he?

* - Soldiers aren't expected to desert, are they?

FOR THE PURPOSES OF CODING, PROCEDURAL AND RHETORICAL QUESTIONS WILL BE IGNORED.

*Occasionally, teachers may use this form of words in phrasing a question, but may nevertheless, in the coder's judgement, have the intention of eliciting a response. (Appropriate clues may be the inflexion of the voice or a pause in the apparent expectation of response.) If this occurs, the question should be coded according to its type (Higher, Middle or Lower Order). A question, in whatever form, should be judged as rhetorical only if there is no apparent intention of eliciting a response.

5. General Coding Rules relating to Questions

(a) An attempt must be made by the coder to note each teacher question (other than procedural and rhetorical questions) under its appropriate category.

(b) In deciding on the category of a question (Higher, Middle or Lower), the criterion should be the apparent intention of the teacher in asking the question, as discernible from the form of words in which the question is expressed and the development of the lesson prior to the question, in so far as this provides evidence of the pupils' level of information, degree of understanding, etc.

The teacher's intention will, of course, be influenced by his knowledge of or assumptions about the amount of relevant data which the pupils possess. Such data may well have been acquired prior to the lesson and the teacher may be aware of this. The coder, however, has no means, beyond the evidence of the lesson itself, of sharing the teacher's knowledge and assumptions, or of determining how much the pupils already know (an amount which may, in any case, vary from pupil to pupil). He must therefore accept the form of the question as his guide to the teacher's apparent intention, unless the evidence of the lesson prior to the question points clearly in another direction.

To give an example: a question in the form "Why did the Corn Laws provoke so much discontent in Britain?" would indicate that the teacher's intention was to analyse the relevant factors. It would therefore normally be categorized as Higher Order. If, however, in an earlier part of the lesson, the teacher had ascertained that the pupils had already received instruction on the/

the Corn Laws and their effects, his question "Why did the Corn Laws---?" would appear to require no more than straightforward recall (or, at the most, comprehension) of information previously acquired.

(c) In deciding on the category of a teacher question, it is important to apply the criteria in (b) above and not to accept question stems as an automatic guide to categories. Examples of question stems that can apply to several categories are

(i) "Give me examples of -----"

e.g. - Give me examples of the things Romans built in Britain.

This is a Lower Order question, demanding recall of information.

- From your knowledge of mediaeval warfare, give me examples of the kind of weapons you would have found at the battle of Stirling Bridge.

This is a Middle Order question, demanding application of generalised information to a specific problem.

- From your knowledge of John Knox and his beliefs, give me examples of the kind of things he might have had to say about Mary's attitude to the Catholics.

This is a Higher Order question, demanding an original communication.

(ii) "What would X have done-----?"

"What would be likely to happen if-----?"

"What would you suppose-----?"

"What would Y have used-----?"

"Where would Z have been found-----?"

"When would they have done this-----?"

"Who would have been responsible for-----?" etc.

History teaching is characterised by a great many "exploratory" questions of this type, requiring pupils to offer suggestions. In terms of the level of cognitive activity involved in the response, such questions appear to lie along a continuum, and the difficulty for the coder is therefore to decide upon the cut-off points between Lower, Middle and Higher Order categories.

The content and context as well as the form of such questions must be considered in judging the apparent intention of the teacher. Compare, for example, the following two questions:

(a) - What change in education policy would occur if there were to be a change in government?

Here the pupils are being asked to make a prediction from the consideration of all the evidence at their disposal. They are required to discern relationships among many elements of information and, through a logical development of consequences, to produce what is, for them, a newly created pattern; in other words, to engage in an act of synthesis.

(b) - What changes in the sound of a recording would occur if the tape replay speed were to be changed from $3\frac{3}{4}$ i.p.s. to $7\frac{1}{2}$ i.p.s.?

The form of question (b) is the same as that of question (a); but the response would appear to involve no more than the direct application of general principles to a specific problem. Note that question (a) is open-ended, whereas question (b) can have only one correct response.

In categorizing this type of exploratory question, the coder must ask himself

- does the question appear to demand no more than recall of information/

information or random guesswork?

e.g. - What furniture would you find in a Roman villa?

can be simply another way of asking

- Do you happen to know what furniture there was in a Roman villa?

Such questions should be coded as Lower Order.

- does the question appear to demand no more than the direct extension of trends apparent in the material being studied or the direct application of principles previously learnt to a specific situation?

e.g. - If trade continued to expand at the rate shown in the table, what would be the probable trade figure for 1980?

- When would the Vikings have used sails and when would they have used oars?

Such questions should be coded as Middle Order.

- does the question appear to demand a fresh act of analysis/ synthesis, involving a perception of relationships and a development of consequences, and leading to a new product?

e.g. - If the Germans had not been halted at El Alamein, what do you think the effect would have been on the course of World War II?

Such questions should be coded as Higher Order.

(iii) Questions with "why?" stems.

Questions in this form normally indicate a requirement for analysis and should therefore be coded as Higher Order. Occasionally, however, the form and/or context of the question indicates a lower category.

e.g./

e.g. - In your document it says that "this weekend excursion owed its appeal to its manifest economy and the salubrious effects of the rural environment." Can you explain in your own words why it was popular?

This is a Middle Order question (see 4B(iv)).

e.g. - Does anyone know why Chamberlain flew to Munich in 1938?

- Can you remember why James IV decided to go to war against the English?

- Why do you think the Boers disliked the British? Have a guess.

These are all Lower Order questions (see 4A(i) and (v)).

(iv) "Do you think-----?"

"What do you think-----?"

Questions beginning "Do you think-----?" often appear to invite no more than a simple yes/no response, but this does not necessarily relegate them to the Lower Order category. The coder must judge the teacher's intention from the context amd content of the questions. For example,

- Do you think you would be frightened if a mediaeval knight came charging at you with his lance?

appears to demand no more than an immediate, affective reaction.

But

- Do you think that an archaeologist is morally justified in breaking into a tomb for the sake of discovering historical evidence?

would appear to involve a Higher Order level of judgement, even though it calls for no more than a yes/no response.

The/

The same considerations apply to "What do you think-----?" questions. Compare, for example,

- What do you think of that story? (Lower Order)

with - What do you think the likely outcome would be if there were to be a general election this month? (Higher Order)

Each question must be judged on its own merits of context and content. The coder should ask himself whether the question demands

- an unconsidered reaction (4A(iv))
- the direct application of principles, etc. to a new situation (4B(vii))
- a fresh act of analysis/synthesis (4C(vii)-(ix))

(v) "Do you mean-----?"

"What do you mean-----?"

These stems are associated with Probing questions (see 6(b)(ii) and (iv)). In determining the appropriate category, the guide must be the level of cognitive work involved for the pupil - recall, comprehension or analysis.

(vi) "Do you agree/disagree-----?"

"Who agrees/disagrees-----?"

Questions in this form are associated with the Redirection of initial responses (see 6(c)). They should be coded in the same category as the question which has produced the response to which they refer, since it must be assumed that reaching agreement or disagreement with a response involves the same order of thinking as that required in making the response.

(d)/

(d) Should the coder be unable to decide on the category appropriate to a question, he should opt for the lower (or lowest) of the categories that he has in mind.

(e) A "pause for response" is defined as a pause of one second or more following a "teacher question", before the teacher resumes speaking. The pause appears to anticipate a pupil response (see 5(j)(i) and (ii)).

(f) Each repetition (whether verbal or implicit) of a question should be coded as a separate question. Such repetition can occur in various ways, for example

(i) where the teacher directs the same question to several pupils in succession

e.g. - T: What kind of things did the Romans build in Britain? John?

P1: Roads.

T: Yes, roads. Anything else? Mary?

P2: Canals?

T: No, I don't think so. Angus?

P3: (no response)

T: Come on, now. Let's have some more suggestions. Jean?

P4:Villas-----etc.

This sequence should be coded as four separate Lower Order questions.

(ii) where the teacher asks a question (probably open-ended), discusses the response and then asks the same question again.

e.g./

e.g. - T: What benefits do you think the Romans brought to Britain?

P1: They made the country more peaceful.

T: Yes. (Discusses inter-tribal warfare and the implications of the Pax Romana). Now, what other benefits do you think the Romans brought?

This sequence should be coded as two separate questions (probably Middle Order).

(iii) where the teacher asks a question and, without giving an opportunity for response, repeats the question or asks another question, perhaps after an intervening comment.

e.g. - T: When did the Romans first come to Britain? When did they first come to Britain?

T: Whereabouts in this district can you find traces of the Antonine Wall? You may be surprised at this, but it's right on your doorstep, so to speak. Which of you has visited the cemeterynear the Bearsden municipal buildings?

These sequences should be coded as two separate Lower Order questions, the first with no opportunity given for response. This rule applies even when the two questions are linked by a conjunction such as 'and' or 'or'.

e.g. - T: Would you prefer to sail in a modern ship and why might it be better?

T: Would there have been any difference in the size of wooden ships, or could they have built them any size they wanted?

Both/

Both these sequences should be coded Qn Q.

(iv) where a teacher asks a question, pauses for a response which is not forthcoming and then rephrases the question in a simpler form, as a means of prompting the pupils.

e.g. - T: What was the principal means of rural transport in the 18th century? (Pause - no response). How did the country people move around from village to village two hundred years ago?

This sequence should be coded as two separate Lower Order questions, the second one being a Prompt (see 6(a) below).

(g) When a question is preceded or followed by a relevant comment, without an intervening pause for response, the comment should be regarded as part of the question.

e.g. - T: Put yourself in the place of Wallace. What do you think he would have done if he'd lost the battle of Stirling Bridge?

T: What do you think Wallace would have done if he'd lost the battle of Stirling Bridge? Put yourself in his place.

n.b. (i) If the question is followed by a pause for response, the subsequent comment should be regarded as a separate entity and should be considered as a prompt (see 6(a))

e.g./

e.g. - T: What do you think Wallace would have done if he'd lost the battle of Stirling Bridge? (Pause) Put yourself in his place (Pause) Yes?

This sequence should be coded Qx P.

(ii) If, however, a question is followed by an instruction such as "Hands up if you know the answer", "Think carefully before you reply", "Take your time and then give me a complete sentence", etc., the instruction(s) should be regarded as part of the preceding question, whether or not there is a pause.

e.g. - T: What kind of man was Machiavelli? (Pause) Think carefully before you answer. (Pause) Now, let's see a show of hands. (Pause) Yes, James?

(h) **Irrespective of whether a question is followed by a pause, the first appeal for a response (whether to the class in general or to a specific pupil) should be regarded as part of the first asking of the question.**

e.g. T: Who succeeded James IV to the throne of Scotland? Does anyone know?

T: Why did Elizabeth mistrust Mary Queen of Scots? (Pause) John, can you tell me?

Sequences which include subsequent appeals for response are to be coded as follows:

- question + general appeal + specific nomination

- CODE as one asking of question

e.g. T. Who built the Antonine Wall? Can anyone tell me? (Pause) Ian?

- question + general appeal + renewed general appeal

- CODE renewed general appeal as second asking of question

e.g. T. Why did Cromwell destroy so many churches? Does anyone have any ideas? (Pause) Come on, it's not difficult

n.b. If the second appeal follows the first without pause for response, code the first asking as "no opportunity given to respond". If a pause for response occurs between first and second appeals, code the first asking as "no response received". See section 7.

- question + specific nomination + general appeal

- CODE general appeal as second asking of question

e.g. T. What is the meaning of "tolbooth"? Jessie? (Pause - no response) Anyone?

- question + specific nomination + subsequent specific nominations

- CODE subsequent nominations as separate askings of the question

e.g. T. Give me the date of the battle/

battle of Bannockburn.
Willie? (Pause - no response)
Mary? (P. "I don't know")
Angus? etc.

"The use of an interrogative "No? No one? Nobody?" etc. should be noted in this context.

e.g. T. Can anyone tell me why Elizabeth distrusted Mary?

(pause for response) No? (or Nobody? etc.)

(pause for response) Jean, you have a try.

In this instance, the No? should be coded as a second asking of the initial question. If, however, no pause for response is allowed after the No? etc., then the word should be regarded as rhetorical (see 4D(b)) and should not be coded.

e.g. T. Can anyone tell me why Elizabeth sistrusted Mary?

(pause for response) Nobody? Well, I'll explain.

(i) A question should be regarded and coded as a Yes/No answer question if it appears to demand no more than the simple answer "yes" or "no".

e.g. - Did you like that story?

- Did the nobles revolt against King John?

- Was Elizabeth the First the last of the Tudor monarchs?

n.b./

n.b. Even if questions of this type attract an "I don't know" response, they should be coded as yes/no answer questions."

General appeal questions seeking to establish the pupils' range of knowledge, experience, reactions, etc. (see 4A(iii) and (iv)) should be coded as yes/no answer questions if the same question addressed to a specific pupil would demand no more than a yes/no response.

e.g. - How many of you come from the other side of the river? (addressed to a specific pupil, this would become "Do you come from the other side of the river?")

- Have any of you ever seen it? (addressed to a specific pupil, this would become "Have you ever seen it?")

- Hands up those of you who have heard of the Chartists. (addressed to a specific pupil, this would become "Have you heard of the Chartists?")

- Who thinks Wallace was a hero? (addressed to a specific pupil, this would become "Do you think Wallace was a hero?")

(j) The coder should distinguish between

(i) questions to which no answer is obtained, even though the teacher pauses to allow the pupils to respond. Such questions include those which attract negative responses indicating inability to provide an adequate answer.

e.g./

e.g. - T. John, can you tell me the date of the battle of Flodden?

P. No.

- T. Mary, why do you think James IV decided to go to war against the English?

P. I don't know.

n.b. The coder should exclude from this category questions beginning "Hands up..." or "Does anyone disagree...?", together with general appeal questions in category 4A(iii) (e.g. "Has anyone read...?"). In these cases, pupil silence should be taken as implying a response.

(ii) questions where the teacher gives the pupil no opportunity to respond. This happens when the teacher

- immediately *follows up his question with a statement or another question

- immediately *answers his own question (having presumably decided, even as he was asking the question, that it was too difficult for his pupils).

*i.e. making a pause of less than one second.

- pauses for more than one second, but ignores audible efforts of the/

the pupils to respond (e.g. calls of "sir! sir!") before continuing with a statement or another question.

Questions of type (i) and (ii) should each be noted by the coder (see 7, Coding Symbols).

6. Prompts, Probes and Redirections

(a) Prompting

A prompt is a word or group of words employed by the teacher as a clue to an initial question that has attracted no response (including negative responses that indicate inability to provide an answer) or that has attracted a response which is rejected in its totality by the teacher.

The prompt may consist of

- restructuring the initial question, usually in a simpler form
- providing a hint or clue

e.g. (i) T. What facilities did the Romans have to transport their troops? (Initial question) (Pause - no response)

T. Come on, how did the Romans move their troops from one place to another? (Prompt - restructuring of question)

(ii) T. What was the mock title given to the English king Edward I? (Initial question)

(Pause - no response)

T. It has something to do with dealing blows...(Prompt - providing a hint)

P. The Hammer of the Scots.

(iii) T. What facilities did the Romans have to transport their troops? (Initial question)

P. They sent them from the big towns out to forts in the country.

T. No, you haven't answered my question. How did the Romans move their troops? (Prompt)

P. They moved them by road.

(iv) T. What was the date of Bleriot's first channel crossing? (Initial question)

P. 1905.

T. No. It was just five years before the start of the Great War. (Prompt)

P. 1909.

n.b. If the teacher repeats the question in its initial form or simply invites the pupil(s) to try again, no prompt should be recorded.

The coder should attempt to note each example of prompting when he makes his analysis of the lesson sequence. It should be noted

(i) that questions coded as prompts must always be related in some supportive way to an immediately preceding initial question. The function of a prompt is to lead the pupil(s) by an easier route to achieve the criterion response demanded by the initial question.

(ii) that a question should be coded as a prompt only if an opportunity to respond to the initial question has been given. Thus, if a teacher asks a question and immediately rephrases it or provides a hint without giving time for response to the initial question, no prompt should be recorded.

(iii) that a prompt should always be coded in the same/

same category as the initial question.

(iv) that occasionally a teacher may provide more than one supportive question or statement as a follow-up to an initial question that has attracted no response.

e.g. a) without pausing between follow-up questions

T. How did the Romans move their troops?

(Initial question)

(Pause - no response)

T. Well, did they move them by road or sea? Did they march or sail in ships? (Prompts)

P. They marched.

In this instance, only the first restructured question should be coded as a prompt (see note (ii) above). The coding sequence would be Qx Pn Q (see section 7 below).

e.g. b) pausing for response between the follow-up questions

T. How did the Romans move their troops? (Initial question)

(Pause - no response)

T. Well, did they move them by road or sea? (Prompt)

(Pause - no response)

T. Did they march or sail in ships? (Prompt)

P. They marched.

In this instance, the second follow-up question is taken as a prompt to the first follow-up question. The coding sequence would therefore be Qx Px P.

e.g. c) without pausing between follow-up statements

T. Who was the German leader during the Second World War? (Initial question)

(Pause - no response)

T. He was a small man with dark hair. His name began with an H. (Prompts)

In this instance, the two statements should be regarded as forming a single prompt. The coding sequence would be Qx P. This applies even if the statements are followed without pause by a repetition of the question, in its original form or in a restructured form (see 5(g)).

T. He was a small man with dark hair. His name began with an H. Does anyone know it?

This sequence should be regarded as forming a single prompt.

e.g. d) pausing for response between follow-up statements

T./

T. Who was the German leader during the second World War? (Initial question)

(Pause - no response)

T. He was a small man with dark hair... (Prompt)

(Pause - no response)

T. His name began with an H...

P. Hitler.

In this instance, a pause after the first prompt indicates to the pupils that a response is still expected. The sequence should therefore be coded Ox Px P.

(v) that if a teacher directs a question to pupil A, obtains no response or rejects the response that is offered, and thereafter directs a restructured question or offers a hint to pupil B, without giving pupil A a chance to respond, no prompt should be recorded.

(b) Probing

Probes are follow-up questions relating directly to a pupil's response and requiring the pupil to develop that response in one of the following ways:

(i)/

(i) by improving the accuracy of a response which the teacher indicates to be broadly acceptable

e.g. P. (replying to an initial question) A centurion was a kind of non-commissioned officer in charge of a company of a hundred men.

T. That's roughly correct, but it was rather less than a hundred men. Can you remember the figure I gave you? (Probe)

P. Oh yes, between eighty and ninety.

(ii) by improving the clarity, through elaboration, completion or rephrasing in a more coherent form

e.g. T. Can you explain some of the factors that brought the first World War to an end in 1918?

P. Tanks.

T. Can you be a bit clearer about what you mean? (Probe)

P. I think tanks were the most important reason for the war ending then.

T. In what way were they important? (Probe)

P. They could over-run the trenches, so once they were invented it meant the end of the kind of trench fighting that had been going on for years.

(iii) by justifying or supporting an assertion or opinion embodied in the response

e.g. T. Who would you say was chiefly to blame for the Peterloo Massacre?

P. Henry Hunt.

T. Why should you single him out? (Probe)

P. Because if he hadn't led them on there would never have been a demonstration in the first place.

(iv)/

(iv) by making the response more precise or specific

e.g. T. What would they eat?

P. Animals.

T. What kind of animals, do you think? (Probe)

P. Deer.

The coder should attempt to note each example of probing when he makes his analysis of the lesson sequence. It should be noted

(i) that questions should be coded as probes only if they are assumed to be directed to the pupil who has made the response which is being probed. This assumption should always be made unless the teacher specifically invites one pupil to answer the initial question and another pupil to answer the follow-up question.

This situation can arise only

(a) through nomination of different pupils

e.g. T. Ian, who benefited most from the Reform Act?

P. The working classes.

T. That's interesting. Jack, can you suggest why?

(b) through the use of a phrase at the follow-up stage which specifically excludes the pupil who responded to the initial question

e.g. T. Who benefited most from the Reform Act?

P. The working classes.

T. That's interesting. Can anyone else suggest why?

Each/

Each of the above sequences should be coded Q Q.

(ii) that a distinction should be made between probing questions, which always relate back to an initial response and are designed to make the pupil improve that response in some way; and linking questions, which are designed to lead the pupil on from his initial response to a fresh aspect of the topic.

e.g. T. Who benefited most from the Reform Act?

P. The working classes.

T. If that is so, what do you think the attitude of the property owners would have been?

In this example, the second question is not a probe. It does not require the pupil to improve on his first response; it leads him on to a related but separate issue.

(iii) that a probing question refers only to responses which the teacher judges to be capable of development. A response which is rejected in its totality cannot be probed, though the teacher may follow up the rejection by prompting the pupil, in order to encourage an accepted response.

(c) Redirecting

For the purpose of this study, redirecting is defined as the redirection of pupil responses to other pupils or to the class in general, as a means of inviting comment and stimulating discussion. "Redirection", as used in this study, does not refer to the redirection of a teacher question to several pupils in succession.

Examples of redirection are

(i) P. (in response to a question) I think Henry Hunt was to blame for the Peterloo Massacre.

T. John, would you agree with that? What do you feel yourself? (Redirection)

(ii) T. Do you think the Antonine Wall was a good way of holding back the Scottish tribes?

P. I think it was pretty useless.

T. Who'd take a different view? (Redirection)

(iii) T. Why do you think so many young children were employed in 19th century factories?

P. Because there were a lot of simple jobs that children could easily do.

T. Mary, would you agree about that as the main reason or would you add anything? (Redirection)

(iv) T. If you were an archer, which tactic do you think you would adopt? Keeping out of range or trying to get in close? Ian?

P1. Keeping out of range.

T. Now why does he say that? Alastair? (Redirection)

P2. There's an awful lot of archers compared with the amount of spearmen.

The coder should attempt to note each example of redirection when he makes his analysis. The following points should be observed:

(i)/

(i) To be classified as a redirection, a question must make reference to and invite comment upon a preceding response (though not necessarily the immediately preceding response). Thus, in example (ii), the teacher's question "Who'd take a different view?" clearly refers to and invites comment upon the pupil's response "I think it was pretty useless." If the teacher said simply, "Any other answers?" his question should be coded as a second asking of the initial question. No redirection would be involved, since "Any other answers?" does not refer to or invite comment upon the initial response.

Similarly, a teacher question, following a response, in the form "Can anyone add anything to that?" should not be coded as a redirection, since, although it follows up the response, it does not invite the pupils to comment - it merely calls for supplementary information, suggestions, etc. (Thus the teacher's follow-up questions in example (iii) would be coded Ryn Q). If, however, the follow-up question is in the form "Would anyone like to add a comment on that?" it should be coded as a redirection.

Note, too, the sequence

T. What was the name given to the Parliamentary soldiers in the 17th century?

P1. Skinheads.

T. Certainly not. Angus?

P2. Roundheads.

Here the teacher is not inviting comment upon the initial response. He makes a comment himself and then repeats his question. No redirection is involved.

This example should be contrasted with the following sequence/

sequence:

T. Philip didn't try to follow Edward over the river. In view of what we have said, why do you think not?

P. Because he couldn't find the crossing.

T. I'm not sure that's the right answer. Why not?

Here the teacher makes a comment upon the initial response, but proceeds to invite further comment from the pupils. The question should therefore be coded as a redirection.

(ii) A redirection sometimes takes the form of several brief questions aimed one after another at the same pupil or at the class in general. Example (i) is a case in point. If this occurs, code each question separately. Thus, example (i) would be coded Ryn R.

(iii) Note that in example (iv) the "why?" question would have been coded as a probe if it had been directed to the pupil who made the initial response. However, it is directed to another pupil, inviting him to comment upon the initial response, and should therefore be coded as a redirection. (See also 6(b), note (i)).

7. Coding symbols

The following symbols should be used in completing the coding sheets:

Q	question	Higher Order, Middle Order, Lower Order, as appropriate
P	prompting question	
Pr	probing question	
R	redirection	

-y yes/no answer question

-x no answer obtained, after pause for response

-n no opportunity given to respond

(i) Symbols y, x, n to be placed, wherever appropriate, after symbols Q, P, Pr, R.

(ii) Symbols x or n will always be placed after y, in cases where the teacher asks a yes/no question and either obtains no response or gives no opportunity to respond.

8. Analysis of sample lesson

The lesson was concerned with the Roman invasion of North Britain, and the teacher's aim was that the pupils should appreciate the importance for the historian of making an objective assessment of two opposing points of view. At the beginning of the lesson, the teacher distributed an extract from Tacitus quoting speeches by the Roman general Agricola and the Celtic leader Calgacus. The speeches ran as follows:

a) Soldiers, you have campaigned with me now for two years already; one spent fighting the tribes near Mona, the other bringing the Roman peace to these northern parts of Britain. Now only one threat to the safety of our province remains: these savage barbarians who live in the hills you see ahead. Now we are all going north to teach these tribes a lesson, that they cannot invade the province of the Roman Emperor with impunity.

(Agricola)

b) We, the last men on earth, the last of the free, have been shielded till today by our very remoteness. But today the boundary of Britain is exposed. Beyond us lies no nation, nothing but waves and rocks, and the Romans more deadly still; the brigands of the world who have exhausted the land by plundering and now they ransack the sea. They create desolation and call it peace. Let us then, unconquered as we are, ready to fight for freedom, prove what heroes Caledonia has been holding in reserve.

(Calgacus)

The teacher then moved on to the questioning and discussion session coded below.

Transcript	HO	MO	LO	Comment	Coding Ref.
T. Now, can anyone see a difference in the point of view expressed in these passages? Stuart?		Q		Providing explanation of material being studied. (No analysis involved)	4B(iv)
P. Sir, one's for the conquering and the other one's against it.					
T. One's for who?				Procedural question. Do not code.	4D(a)
P. One's for conquering, for making everything desolate, they want to get more further on.					
T. Can anyone enlarge on that? What about you?		Q		Explanation of material studied. Not a Probe, since question excludes pupil who made the initial response.	4B(iv) 5(h)
P. Please sir, one's biassed against the tribes of northern Scotland, and the other one's biassed against the Romans.					6(b), note (1) (b)
T. That's right. In other words, one of them is sympathetic/					

Transcript	HO	MO	LO	Comment	Coding Ref.
sympathetic towards the native Celts, the other is sympathetic towards the Roman point of view, the way the history books tend to be. Does this tell you anything about the type of person who would have spoken these words? Yes?		Q		Applying information (n.b. considered for y coding, but reasonable to assume that teacher was expecting more than a yes/no response)	4B(vii)
P. Sir, one's a Roman and one's a Caledonian.					5(i)
T. That's right, one's a Roman general. Agricola's a Roman general, and Calgacus is a leader of the Celts. So this is a bit different from normal - usually you just get the Roman point of view. But in this instance you've got a Roman writer - he was called Tacitus - and he's recording the views of both sides. So how would you be/					

Transcript	HO	MO	LO	Comment	Coding Ref.
be objective here? We've talked about objectivity in a football match. How would you try to be objective here? (Pause) Has anyone got any ideas, to be fair, in other words, without bias? Yes?	Qn Q			Problem solving, involving analysis. No opportunity for response Problem solving First appeal for a response	4C(vii) 5(j)(ii) 4C(vii) 5(h)
P. You could say that the Romans were good conquerors and that the Caledonians tried very hard to defend their country.					
T. Yes, but having both these passages here, how would you try to arrive at a balanced point of view? Does anyone - yes, at the back there.	Q			Problem solving (Not a Prompt, since previous response not totally rejected)	4C(vii)
P. Sir, you could say that the Caledonians were fighting/					

Transcript	HO	MO	LO	Comment	Coding Ref.
fighting to defend their land and that the Romans were just fighting to try to conquer it, because they were frightened that the Scots one day would be strong, and would try to conquer them, so they thought their action would be justified.					
T. Yes, that's true. But what I'm really driving at, as in the football match, you might be biassed towards your team, but in order to be fair, to be objective, you would attempt to assess how good the other team was as well. You would say good points about both teams, bad points about both teams. So to be objective here you'd have to read both the passages and then try to balance the viewpoints, arrive at a fair viewpoint, because/					

Transcript	HO	MO	LO	Comment	Coding Ref.
because obviously both the men speaking have got their own viewpoints and they're sticking to them. Right then, what does the Roman general, that's Agricola, in the first question, see as the task of his soldiers in Britain? It tells you in the passage.				Relevant comment, regarded as part of the question	5(g)
Yes?		Q		Providing explanation of material studied	4B(iv)
P. He says that his soldiers must defeat the Celts in the highlands.					
T. Yes, and do what? It says more than that, does anyone see any more than that?		Prn		Probe for precision, with no opportunity given for response	6(b)(iv) 5(j)(ii)
Yes?		Q		Providing explanation (No longer a Probe, since question directed away from initial responder)	4B(iv)
P. And teach the Celts a lesson.					6(b) note (i) (b)
T. Yes, but there's more than that. The boy at the front here.		Q		Providing explanation	4B(iv)
P./					

Transcript	HO	MO	LO	Comment	Coding Ref.
P. Sir, for the safety of their own province. To keep Rome safe.					
T. Yes, that is true, but there's a specific sentence I'm looking for there. Yes?		Q		Perception of content	4B (intro.)
P. It means that they've got to fight with the Celts and they've got to make sure of the safety of their province at the same time.					
T. Yes, all that's true, but what about the Romans bringing to Britain? What did they think they were bringing to Britain?		Qn		Perception of content	4B (intro.)
				No opportunity for response	5(j)(ii)
		Q		Perception of content	4B (intro.)
P. They would bring peace to Britain.					
T. Yes, that's the exact sentence/					

Transcript	HO	MO	LO	Comment	Coding Ref.
sentence I'm looking for. Yes, there it's there in the third line. "Bringing the Roman peace to Britain." And obviously they thought this was their task in coming to Britain. They were doing the Britons a favour in this way. But what does he think of the native Britons, the Celts? Yes, the boy at the back. (Pause) Does he praise them?		Qx Py		Providing explanation No response after pause Prompt (providing a hint)	4B(iv) 5(e) 5(j)(i) 6(a) + note (iii)
P. No, sir. He thinks that they're plunderers and savages.				Yes/no answer question (even though pupil proceeds to elaborate his response)	5(i)
T. That's exactly it. Where are the exact two words I'm looking for? This boy.			Q	Straightforward recognition of material (the word "savage" had been given in previous response).	4A(i)
P. Sir, "savage barbarians".					
T. Savage barbarians. That's right/					

Transcript	HO	MO	LO	Comment	Coding Ref.
right. So obviously he's not being very complementary to the native Celts who live in Scotland. So this is a Roman point of view quite obviously. Do you think it's fair? (Pause) Yes, this boy here.			Qy	Inviting unconsidered reaction (yes/no answer)	4A(iv), 5(c)(iv) 5(i)
P. No, sir. Please, sir, may-be even if the Romans hadn't come to Britain, they may even still have become civilised.					
T. Yes, that's exactly what I'm looking for. And we'll see this in the next passage. The second passage is obviously exactly the opposite. You're looking at it from the point of view of Calgacus, and he is a leader of the Celts. And at exactly the same time, this is him talking to his men about the Roman soldiers/					

Transcript	HO	MO	LO	Comment	Coding Ref.
soldiers. So the Romans, as we said earlier, think that they're bringing peace to Scotland. Does he think that as well? Yes?		Qy		Stating the main idea in material being studied	4B(v)
P. Sir, he thinks they're bringing death and war.					
T. Can you see the exact part in the passage? That's what I'm looking for. (Pause) Can anyone see the bit there? (Pause) The boy at the back there on the left		Qx Q		Perception of content (no response) Renewed general appeal + specific nomination	4B (intro.) 5(j)(i) 5(h)
P. Sir, "they create desolation and call it peace."					
T. That's it exactly. So that's completely different. In the earlier passage, the Roman general thought he was bringing the Roman/					

Transcript	HO	MO	LO	Comment	Coding Ref.
Roman peace; but to the Celts they're creating desolation and they call that peace. So - and we noticed earlier on that the Romans thought the Celts were savage barbarians. What did they think of the Romans? (Pause) Anyone over in this direction?		Q		Providing explanation	4B(iv)
P. They thought they were plunderers, and they were spoiling the land and things, and the sea.					
T. That's it exactly. And can you see the exact part, anyone, where it says that?			Qx	Recognition of material	4A(i)
(Pause) What does it call the Romans? (Pause)				(no response)	5(j)(i)
The Romans called the Celts savage barbarians, so the Celts - what's their opinion? Yes.			Px	Prompt (providing hint)	6(a)
				(no response)	5(j)(i)
			P	Prompt (restructuring question)	6(a)
P./					

Transcript	HO	MO	LO	Comment	Coding Ref.
P. "The brigands of the world".					
T. That's it exactly. The robbers of the world, the plunderers of the world. That's what I'm looking for. So the Celts, they've been protected in the north of Scotland and it tells you that in the passage. Why were they protected, do you see that?		Q		Providing explanation (no analysis involved, since reason clearly stated in passage)	4B(iv)
P. Among the rocks and waves.					
T. That's right, because they were away up in the hills, far away. So there we have a completely opposed point of view of the same situation. Does anyone think that was a fair point of view, the Celtic point of view? (Pause) Yes, at the back.			Qy	Inviting unconsidered reaction. Yes/no question in the form of general appeal followed by specific nomination. (In fact, the pupil gives a reasoned response, but the coding guide must be the form of the question)	4A(iv) 5(e)(iv) 5(i) 5(h)
P. Sir, it's fair to the same extent/					5(b)

Transcript	HO	MO	LO	Comment	Coding Ref.
extent as the first one was, so it's one-sided.					
T. That's it exactly, exactly. That's it biassed again. It's every bit as biassed as the first one, every bit as biassed as you being a football supporter, just talking about your team and never giving any credit to the other team, no matter how well it plays. So, in order to reach an objective conclusion, what do we have to do? Yes?		Q		Applying principle already acquired	4B(vii)
P. We have to look at the points of view of both sides.					
T. That's it exactly. To reach an objective conclusion, you've got to try to eliminate this subjective side, try to eliminate the bias and balance the two points/					

Transcript	HO	MO	LO	Comment	Coding Ref.
points of view. So is there anyone here agrees with the Romans?			Qyn	Expressing opinion (yes/no question, no opportunity for response	4A(iv) 5(i), 5(j)(ii)
Or anyone with the Celts? (Pause)			Qyx	Expressing opinion (no response)	4A(iv) 5(i), 5(j)i)
No one feels strongly either way?				Interrogative "no one" with no ensuing pause. Rhetorical	5(h) 4D(b)
Who do you agree with?			Q	Expressing opiniom	4A(iv)
P. Sir, I agree with the Celts, because they're liable to be attacked, sir, and they naturally think that they're coming to get all the riches out of the land and that.					
T.Yes. And is there anyone who'd like to argue about that - anyone pro-Roman? Yes, the boy at the back there	R			Redirection of response (which had stated a reasoned argument)	6(c)
P. Please, sir, I'm not pro-Roman, but the Romans brought no civolisation, please, sir, they made bath-houses, please sir, they/					

Transcript	HO	MO	LO	Comment	Coding Ref.
they didn't all want wars, please, sir, they wouldn't have let us come to the towns if they'd wanted war, but they let the Britons come to the towns, they let them go to the baths and they let them go to the theatre.					
T. Mmm.					
P. They built their houses for them, they gave them a good job and a good life.					
T. So they brought civilisation to Britain. What would you say to that? He says they built houses, they built roads, they brought civilisation...	R			Relevant comment, regarded as part of the question Redirection of response (again, teacher appears to be soliciting reasoned comments)	5(g) 6(c)
P. Sir, at the time the Romans attacked, they didn't know they were going to build theatres and/					

Transcript	HO	MO	LO	Comment	Coding Ref.
and baths for them and that.					
T. So in other words, their motives weren't purely to help the native British. They had their own motives, personal motives -					
P. Sir, but the Celts didn't know they were going to do that at the time -				Unsolicited comment	
T. Hm-mm -					
P. Sir, they didn't know that there'd be anybody build them baths or theatres.				Unsolicited comment	
T. Now, that's true. So you've got this one side, you've got the fact that the Romans brought a system of law, they brought peace despite what it says, be-cause they conquered most people and they maintained the/					

Transcript	HO	MO	LO	Comment	Coding Ref.
the peace and a system of law. Anything else they brought, anything you'd like to add to that? Yes, this boy at the back.		Q		Applying information already acquired. (n.b. teacher appears to be asking pupils to select examples of benefits brought by Romans. Question thus goes beyond simple recall)	4B(vii)
P. They brought ways of building houses and streets and everything like that.					
T. Well, I've already said that. They brought ways of building, new architectural ideas. Anything else? Yes?		Q		Applying information	4B(vii)
P. They brought their laws with them.					
T. Well, yes, the system of law, which maintained the peace, that and force of arms, of course. Anything else? This boy at the back.		Q		Applying information	4B(vii)
P. Sir, they brought religion, Roman/					

Transcript	HO	MO	LO	Comments	Coding Ref.
Roman Catholic religion to Britain.					
T. Ah, not at this point that we're talking about, this is a very early time when the Romans had their own gods. No, the religious side doesn't really enter into it. Right, the boy in front		Q		Applying information	4B(vii)
P. They brought their language to Britain.					
T. That's right, that's a good point, they brought the Latin language.					

APPENDIX P

ASSESSMENT OF RECORDED LESSONS : SUMMER TERM, 1973

STUDENT'S NAME ----------

ASSESSED BY ----------

	Weak	Well below average	Slightly below average	Average	Slightly above average	Well above average	Very good
	E	D	C-	C	C+	B	A
1. Optimum use of "higher order" questions (analysis, synthesis, evaluation)							
2. Restriction of "lower order" questions to optimum number							
3. Pausing to give pupils opportunity to respond adequately							
4. Optimum use of Prompting, to stimulate response							
5. Optimum use of Probing questions, to develop initial responses							
6. Appropriate use of Redirection of initial answers, for comment and discussion							

NOTES

a. In deciding on your rating, think in terms of an objective weak - average - good lesson continuum. Do not "make allowances" for the student.

b. In judging the use of a skill, keep in mind the extent to which the student availed himself of his opportunities to use the skill appropriately, in the context of the lesson.
(All students were asked to plan their lessons to include ten minutes of questioning and discussion and to concentrate on the questioning skills listed above.)

c. "Higher Order" questions are those which require the pupil to engage in a process of reasoning, to identify motives and causes, make inferences, draw conclusions and formulate judgments.

d. "Lower Order" questions involve the pupil in straightforward recall of information or recognition of material; or invite him to express opinions, preferences, etc. based on unconsidered reaction. Many one-word answer and yes/no questions fall into this category. It was recognized that a certain number of such questions can be appropriate, but students were encouraged not to over-indulge in them.

e. Pausing for adequate response

This skill is particularly appropriate in the context of higher order questions, where the pupils need time to think out a good answer. The reverse behaviour (asking a question and following it immediately by another question or comment, without giving pupils a chance to respond) would produce a low rating on this skill.

f. A Prompt is defined as a hint or clue used as a follow-up to an initial question which has attracted no response or a response which the teacher has rejected.

g. A Probe is defined as a follow-up question requiring the pupil to develop his initial response by improving its accuracy, clarity or precision, or by justifying it.

h. A Redirection is defined as a question inviting a pupil (or pupils) to comment upon an initial response. Its aim is to stimulate discussion.

BIBLIOGRAPHY

ACHESON K (1964)
Effects of Feedback from Television Recordings and Three Types of Supervisory Treatment on Selected Teacher Behaviour
Unpublished dictoral dissertation, Stanford University, California

ACHESON K A and TUCKER P E (1971)
"Videotape versus written instruction and videotape versus audiotape feedback in a minicourse on higher cognitive questioning"
Far West Laboratory for Educational Research and Development, University of Oregon, U.S.A.
Report A71 - 18

ALLEN D W, BERLINER D O, McDONALD F J and SOBOL F T (1967)
"A comparison of different modelling procedures in the acquisition of a teaching skill"
Paper presented at AERA meeting, New York

ALLEN D W and CLARK R J (1967)
"Microteaching: its rationale"
High School Journal, No. 51, pp 75-79

ALLEN D W and FORTUNE J C (1966)
"An analysis in microteaching: new procedure in teacher education. Microteaching: a description"
School of Education, Stanford University, U.S.A.
Mimeo

ALLEN D W, FORTUNE J C and COOPER J M (1968)
"The Stanford summer microteaching clinic, 1966", in "Microteaching: a description"
Stanford Teacher Education Program, Stanford University, California

ALLEN D W and RYAN K A (1969)
Microteaching
Addison Wesley. Reading, Massachusetts

ANDERSON C C and HUNKA S M (1963)
"Teacher evaluation: some problems and a proposal"
Harvard Educational Review, vol. 33, pp 74-95

BARNES D/

BARNES D, BRITTON J and ROSEN H (1969)
Language, the Learner and the School
Penguin Books Ltd., Harmondsworth

BAYER E (1972)
Multidimensional Analysis of the Verbal Communication in the Classroom
Unpublished doctoral dissertation, University of Liège
Reported in CHANAN G (ed.) (1973), Towards a Science of Teaching, NFER Publishing Co., Ltd., Windsor, Berks

BELL C (1968)
"A report of an investigation of microteaching in the development of teaching performance in home economics education at Texas Technological College"
Lubbock: School of Home Economics, Texas Technological College

BELLACK A A, KLIEBARD H M, HYMAN R T, SMITH F L Jr. (1966)
The Language of the Classroom
Teachers College Press, Columbia University, New York

BERLINER D C (1969)
"Microteaching and the technical skills approach to teacher training"
Stanford Technical Report No. 8, Stanford University, California

BIDDLE B J and ELLENA W J (1964)
Contemporary Research on Teacher Effectiveness
Holt, Rinehart and Winston, New York

BIERSCHENK B (1972)
Self-Confrontation via Closed Circuit Television in Teacher Training: Results, Implications, and Recommendations
Bulletin No. 37, Department of Educational and Psychological Research, School of Education, Malmo, Sweden

BLOOM B S (ed) (1956)
Taxonomy of Educational Objectives
Book 1 Cognitive Domain
David McKay Co. Inc., U.S.A.

BORG W R, KELLEY M L and LANGER P (1970)
Minicourse 1: Effective Questioning: Elementary Level
Macmillan Educational Services Inc., California

BORG W R, KELLEY M L, LANGER P and GALL M (1970)
The Minicourse: a Microteaching Approach to Teacher Education
Macmillan Educational Services, Inc. Beverly Hills, California

BORGER R and SEABORNE A E M (1966)
The Psychology of Learning
Penguin Books Ltd., Harmondsworth, Middlesex

BORTZ E L (1971)
Electronic Recording media, Microteaching Self-Critique, and the Questioning Behaviour of Student Teachers
Unpublished doctoral dissertation, University of Maryland, U.S.A.
University Microfilms No. 72 - 607

BRASHEAR R M and DAVIS O L (1970)
"The persistence of teaching laboratory effects into student teaching: a comparative study of verbal teaching behaviours and attitudes"
Paper presented at meeting of AERA, 1970.
ED 039 176

BRIGGS L J (1968)
Student Response in Programmed Instruction
National Academy of Sciences, National Research Council. Washington, D.C.

BRITTON R J and LEITH G O M (1971)
"An experimental evaluation of the effects of micro-teaching on teaching performance"
R M Phillips Research Unit, University of Sussex.
Mimeo

BROWN G A (1971)
"Microteaching: innovation in teacher education"
Education for Teaching, No. 86, pp 11-15

BRUSLING C (1974)
Microteaching - A Concept in Development
Almquist and Wiksell International. Stockholm

BUSH R N (1966)
"Microteaching: controlled practice in the training of teachers"
Communication, July 1966, 201-207

CICIRELLI V G (1969) "University supervisors' creative ability and their appraisal of student teachers' classroom performances: an exploratory study" Journal of Ed. Res., Vol. 62, pp 375-381

CLAUS K E (1968) "Effects of cueing during modeling and feedback sessions on learning a teaching skill" Paper presented at annual meeting of American Psychological Association, San Francisco

CLAUS K E (1969) "Effects of modeling and feedback treatments on the development of teachers' questioning skills" Technical Report No. 6. Stanford Center for Research and Development in Teaching, Stanford University, California

COLLINS M (1959) "A follow-up study of some former graduate student teachers" Brit. Journal of Ed. Psych., Vol. 29, No. 3, pp. 187-197

COLTHAM J B and FINES J (1971) Educational Objectives for the Study of History TH 35, The Historical Association, London

COPELAND W D and DOYLE W (1973) "Laboratory skill training and student teacher classroom performance" The Journal of Experimental Education, Vol. 42, No. 1, pp 16-21

COREY S M (1940) "The teachers out-talk the pupils" School Review, 48, 745-752

DAVIS O L and TINSLEY D C (1967) "Cognitive objectives revealed by classroom questions asked by social studies student teachers" Peabody Journal of Education, 1967, 45, 21-26

DE LANDSHEERE G (1970) "Le formation des maitres par l'analyse des interactions pedagogiques" Reported in CHANAN G (ed) (1973), Towards a Science of Teaching, NFER Publishing Co., Ltd., Windsor, Berks.

FLANDERS N A (1963) "Intent, action and feedback: a preparation for teaching" J. Teacher Educ., 14, 251-260

FLANDERS N A (1970) Analysing Teacher Behaviour Addison-Wesley, Reading, Massachusetts

FLOYD W D (1960) An Analysis of the Oral Questioning Activity in Selected Colorado Primary Classrooms Unpublished doctoral dissertation, Colorado State College

FORTUNE J C, COOPER J M and ALLEN D W (1967) "The Stanford summer microteaching clinic, 1965" Journal of Teacher Education, 18, 4, 389-393

FRANCIS K A E (1971) A Study of the Effect of Teacher Intention and Class Level on Teaching Behaviour and Pupil Thinking and Learning in the Classroom Unpublished Master's thesis, University of Canterbury, New Zealand Reported in NUTHALL G A and CHURCH J, "Experimental studies of teaching behaviour"; in CHANAN G (ed), (1973), Towards a Science of Teaching, NFER Publishing Co., Windsor, Berks

FULLER F F and MANNING B A (1973) "Self-confrontation reviewed: a conceptualization for video playback in teacher education" Review of Ed. Res., Vol. 43, No. 4, pp 469-528

GAGE N L (1968) "An analytical approach to research on instructional methods" Journal of Experimental Education, Vol. 37, No. 1

GALL M D (1972) "The use of questions in teaching" Review of Ed. Res., Vol. 40, No. 5

GALL M D, DUNNING B and WEATHERSBY R (1971) Minicourse 9: Higher Cognitive Questioning (Teachers' Handbook) Macmillan Educational Services Inc., California

GALL M D et al (1971) "Improving teachers' mathematics tutoring skills through microteaching: a comparison of videotape and audiotape feedback" Paper presented at the Annual Meeting of the AERA, New York

GALLACHER J J (1965)
"Expressive thought by gifted children in the classroom"
Elementary English, 1965, 42, pp 559-568

GALLACHER J J and ASCHNER M J (1963)
"A preliminary report: analysis of classroom interaction"
Merrill-Palmer Quarterly of Behavior and Development July, 1963, pp. 183-194

GRANT D (1974)
The Case for Microteaching in Teacher Training
Unpublished Master of Education thesis, Department of Education, University of Liverpool

GREGORY I D (1971)
"Microteaching in a pre-service education course for graduates"
British Journal of Educational Technology, No. 1, Vol. 2, pp 24-32

GRIFFITHS R (1972)
"The role of the tutor in microteaching supervision: a survey of research evidence"
University of Stirling

GRIFFITHS R (1973a)
"Ten years of microteaching"
University of Stirling

GRIFFITHS R (1973b)
"The future development of microteaching techniques - some possibilities"
Mimeo. University of Stirling

GRIFFITHS R (1974)
"The training of microteaching supervisors"
Department of Education, University of Stirling
Paper presented at the Conference "Perspectives in Microteaching", University of Liverpool, September, 1974

GUILFORD J P (1956)
"The structure of the intellect"
Psych. Bulletin 53, pp 267-293

HAYNES H C (1935)
The Relation of Teacher Intelligence, Teacher Experience and Type of School to Types of Questions
Unpublished doctoral dissertation, George Peabody College for Teachers

HEAPS/

HEAPS J F (1973)
"Improvements in teacher training techniques in Kenya"
Educational Broadcasting International, Vol. 6, No. 2, pp 73-75

HOUGH J B and DUNCAN J K (1970)
Teaching: Description and Analysis
Addison-Wesley Publishing Co., Reading, Massachusetts

HUGHES M et al (1959)
Development of the Means for the Assessment of the Quality of Teaching in Elementary Schools
University of Utah. Salt Lake City

HUNKINS F P (1967)
"The influence of analysis and evaluation questions on achievement in sixth grade social studies"
Paper presented at the annual meeting of the American Educational Research Association, New York, 1967

HUNKINS F P (1968)
"The effects of analysis and evaluation questions on various levels of achievement"
Paper presented at the annual meeting of the American Educational Research Association, Chicago, 1968

IVEY A E and ROLLIN S A (1974)
"The human relations performance curriculum: a commitment to intentionality"
British Journal of Ed. Tech., No. 2, Vol. 5, pp 21-29

JACOBSEN D A and GERLACH V S (1973)
"A comparison of the effects of textual and televised modes of instruction in teacher education"
University of Georgia, U.S.A.
EDRS ED 074 773

JARDINE R (1972)
"An exploration in the use of videotape recording in teacher-pupil relationships"
Visual Education, March 1972, pp 21-27

JOHNSON J A (1968)
"A national survey of student teaching programs"
Northern Illinois University. Mimeo

JOHNSON W D (1967)
"Microteaching: a medium in which to study teaching"
High School Journal, Vol. 51, pp 86-92

JOHNSON W D and KNAUPP J E (1970)
"Trainee role expectations of the microteaching supervisor"
Journal of Teacher Education, Vol. 21, pp 396-401

KALLENBACH W W and GALL M D (1969)
"Microteaching versus conventional methods in training elementary intern teachers"
Journal of Ed. Res., 63, 3, pp 136-141

KELLY G U (1973)
Microteaching at Coventry College of Education
Research Project HR 2104/1, Social Science Research Council, 1972-73, Coventry, Warwickshire

KERLINGER F N (1964)
Foundations of Behavioural Research
Holt, Rinehart and Winston, New York

KISSOCK C McN (1971)
A Study to Test the Value of Microteaching in a Program of Video Modeling Instruction in the Development of Higher Order Question Asking on the Part of Pre-Service Teachers
Unpublished doctoral dissertation, University of Minnesota. University Microfilms 71-28.252

KLINGSTEDT J L (1970)
Effectiveness of Three Microteaching Feedback Procedures
Unpublished doctoral dissertation, Tech. University, Texas, U.S.A.

KOMISAR N P (1966)
"Conceptual analysis of teaching"
High School Journal, 50, 14-21

KORAN J J (1968)
"The relative effects of classroom instruction and subsequent observational learning on the acquisition of questioning behaviour by pre-service elementary science teachers"
(Science Education Centre)
University of Texas, Austin, U.S.A.

KORAN M L (1969)
"The effects of individual differences on observational learning in the acquisition of a teaching skill"
AERA Conference Report, Los Angeles, U.S.A.

KORAN/

KORAN M L
(1972)
"Varying instructional methods to fit trainee characteristics"
A.V. Communication Review, 20, 2, pp 135-146

KORAN M L, McDONALD F J and SNOW R E
(1969)
"The effects of individual differences on observational learning in the acquisition of a teaching skill"
Paper presented to the annual conference of the AERA Mimeo

KORAN M L, SNOW R E and McDONALD F J
(1971)
"Teacher aptitude and observational learning of a teaching skill"
Journal of Ed. Psych., 62, 1971, pp 219-228

LANGE D N
(1971)
"An application of social learning theory in affecting change in a group of student teachers using video modeling techniques"
The Journal of Educational Research, Vol. 65, No. 4, pp 151-154

LAWLER E S
(1964)
"Differing rates of progress of classes under the same and different teachers"
Journal of Ed. Res., Vol. 58, pp 84-86

LEONARD B C, GIES F J and PADEN J S
(1971)
"The effect of selected media feedback upon the interactive behaviour of student teachers"
The Journal of Ed. Res., Vol. 64, No. 10

LEVIS D et al
(1973)
"A progress report of the study to investigate the effects of alternative techniques to practice teaching on the development of teaching skills by student teachers"
School of Education, Macquarie University, Sydney

McDONALD F J and ALLEN D W
(1967)
"Training effects of feedback and modelling procedures on teacher performance"
Technical Report No. 3, Stanford Center for Research and Development in Teaching, Stanford University, California

McINTYRE D
(1971)
"Three approaches to microteaching: an experimental comparison"
Department of Education, Univrsity of Stirling. Mimeo

MEDLEY D M and MITZEL H E (1959)
"Some behavioural correlates of teacher effectiveness" The Journal of Ed. Psych., Vol. 50, pp 239-246

MEDLEY D M and MITZEL H E (1963)
"Measuring classroom behaviour by systematic observation" In GAGE N L (ed) Handbook of Research on Teaching, Rand McNally, Chicago

MEUX M and SMITH B O (1964)
"Logical dimensions of teaching behaviour" in BIDDLE B J and ELENA W J (eds), Contemporary Research on Teacher Effectiveness, Holt, Rinehart and Winston, New York

MILLAR C J (1972)
A Procedure for Analysing Evaluations of Observed Teaching and its Application in Measuring Outcomes of Professional Education Unpublished MSc thesis, University of Stirling

MORRISON A and McINTYRE D (1973)
Teachers and Teaching (2nd edition) Penguin Education, Harmondsworth, Middlesex

MORSE K R, KYSILKA M L and DAVIS O L (1970)
"Effects of different types of supervisory feedback on teacher candidates' development of refocusing behaviours" Research and Development Centre for Teacher Education, University of Texas, Report Series No. 48

NIAS J (1974)
"Beyond microteaching: other uses of video recording in teacher training" Visual Education, December 1974, pp 23-26

NINANE A M (1969)
Essai de 1'Application de la Taxonomie de Bloom (Domaine Cognitif) a 1'Analyse des Comportements d'Enseignement Unpublished dissertation, University of Liege Reported in CHANAN G (ed) Towards a Science of Teaching, NFER Publishing Co. Ltd., Windsor, Berks

NUTHALL G A and CHURCH J (1973)
"Experimental studies of teaching behaviour" in CHANAN G (ed) Towards a Science of Teaching, NFER Publishing Co, Windsor, Berks

OPPENHEIM/

OPPENHEIM A N (1970) — Questionnaire Design and Attitude Measurement. Heinemann, London

ORME N E J (1966) — Effects of Modelling and Feedback Variables on the Acquisition of a Complex Teaching Strategy. Unpublished doctoral dissertation, Stanford University, California

OWENS L and HATTON N (1970) — "'Telling it like it is' - microteaching in a teacher education program". Mimeo. University of Sydney

PARLETT M and HAMILTON D (1972) — Evaluation as Illumination: a New Approach to the Study of Innovatory Programs. Occasional Paper 9, Centre for Research in the Educational Sciences, University of Edinburgh

PHILLIPS N E (1973) — "Effects of a videotaped modelling procedure in verbal questioning practices of secondary social studies student teachers". Fairmont State College, West Virginia, U.S.A. EDRS ED 079 967

POPPLETON P K (1968) — "The assessment of teaching practice: what criteria do we use?". Education for Teaching, Vol. 75, pp 59-64

ROBERTSON J D C (1957) — "An analysis of the views of supervisors on the attributes of successful graduate student teachers". Brit. Journal of Ed. Psych., Vol. 27, No. 2, pp 115-126

ROSENSHINE B (1970) — "Enthusiastic teaching: a research review". School Review, August 1970, pp 499-514

ROSENSHINE B and FURST N (1971) — "Research on teacher performance criteria" in SMITH B O (ed) Research in Teacher Education: a Symposium. Prentice-Hall Inc. Englewood Cliffs, N J

RUTHERFORD R B (1973) — "The effects of a model videotape and feedback videotapes on the teaching styles of teachers in training". The Journal of Experimental Education, Vol. 42, No. 1, pp 64-69

RYANS/

RYANS D G (1960) — Characteristics of Teachers: their Description, Comparison and Appraisal. American Council on Education. Washington D.C.

SALOMON G and McDONALD F J (1969) — "Pre- and post-test reactions to self-viewing one's teaching performance on videotape" Stanford Center for Research and Development in Teaching. R and D memo No. 44

SANDERS N M (1966) — Classroom Questions - What Kinds? Harper and Row, New York

SHIPLEY C M, CANN M M, HILDEBRAND J and MITCHELL G T (1968) — A Synthesis of Teaching Methods McGraw-Hill, U.S.A.

SHIPMAN M D (1966) — "The assessment of teaching practice" Education for Teaching, Vol. 70, pp 28-31

SHIVELEY J E et al (1970) — "The effect of mode of feedback in microteaching" Paper presented at AERA Conference ED 037 391

SKAILAND D (1972) — "Minicourse 18: main field test report" Far West Laboratory for Educational Research and Development, Berkeley, California. Mimeo

STEVENS R (1912) — "The question as a measure of efficiency in instruction" Teachers College Contributions to Education, No. 48 Teachers College, Columbia University

STONES E and MORRIS S (1972) — Teaching Practice: Problems and Perspectives Methuen and Co. Ltd. London

TABA H and HILL J J (1965) — Teacher Handbook for Contra Costa Social Studies, Grades 1-6 Hayward. California

TAYLOR/

TAYLOR M T and SHARP R M (1971) "Classroom analysis for pre-service teachers" Education for Teaching, No. 86, pp 16-23

TINSLEY D C and DAVIS O L (1969) "The relationship of student teachers' opinions of intellectual emphasis and judged cognitive levels of questions framed for discussion and tests" AERA Paper Abstracts, 1969, pp 108-109

TITTLE C J and HANDLE C (1970) Research on Practical Experience in Teacher Education: a Selected, Annotated Bibliography Report No. 70-12, Office of Institutional Research and Program Evaluation Division of Teacher Education, The City University of New York

TUCKMAN B W and OLIVER W F (1968) "Effectiveness of feedback to teachers as a function of source" Journal of Ed. Psych., Vol. 59, No. 4, pp 297-301

TURNEY C (1970) "Microteaching - a promising innovation in teacher education" The Australian Journal of Education, 14(2), pp 125-141

TURNEY C, CLIFT J C, DUNKIN M J and TRAILL R D (1973) Microteaching: Research, Theory and Practice Sydney University Press. Sydney

UNRUH W R (1968) "The modality and validity of cues to lecture effectiveness" in GAGE N L et al, Explorations of the Teacher's Effectiveness in Explaining: Technical Report No. 4 Stanford Center for Research and Development in Teaching

WAGNER A C (1973) "Changing teacher behaviour: a comparison of micro-teaching and cognitive discrimination training" Journal of Ed. Psych., Vol. 64, No. 3, pp 299-305

WALKER R (1972) "The Sociology of education and life in school classrooms" International Review of Education, XVIII, No. 1

WARD B E (1970)
"A survey of microteaching in NCATE-accredited secondary education programs"
Research and Development Memorandum No. 70, Stanford Center for Research and Development in Teaching, Stanford University

WARD P M (1970)
"The use of the portable videotape recorder in helping teachers self-evaluate their teaching behaviour"
University of California, U.S.A.

WHITE D R (1972)
The Stirling Lesson-Sampling Instruments: the Preparation, Testing and Trial Use of a Battery of Lesson-Sampling Instruments
Unpublished MSc Thesis, University of Stirling

WISEMAN S and START K B (1965)
"A follow-up of teachers five years after completing their training"
Brit. Journal of Ed. Psych., Vol. 35, No. 3, pp 342-361

WOOD C C and HEDLEY R L (1968)
"Training instruction practice sessions (TIPS): observations on student reaction to the use of video-tape recordings (VTR) in simulated classroom situations"
Canadian Education and Research Digest, Vol. 8, No. 1, pp 46-59

WRAGG E C (1973)
"A study of student teachers in the classroom"
in Towards a Science of Teaching CHANAN G (ed), NFER Publishing Co. Ltd., Windsor, Berks

WRIGHT C (1973)
Report on Microteaching Experiment
Hamilton College of Education, Hamilton, Lanarks.
Mimeo

WRIGHT E M J and NUTHALL G A (1970)
Report by NUTHALL G A and CHURCH J in "Experimental studies of teaching behaviour"; in CHANAN G (ed) (1973) Towards a Science of Teaching, NFER Publishing Co. Ltd., Windsor, Berks

YOUNG D A (1970)
"Preliminary report on the effectiveness of supervision on the acquisition of selected teaching behaviours in a microteaching series"
Paper presented at the Annual Meeting of AERA

YOUNG D B (1967) The Effectiveness of Self-Instruction in Teacher Education Using Modelling and Videotape Feedback Unpublished doctoral dissertation, Stanford University, California

---

ARGYLE M (1969) Social Interaction Methuen, London

BORG W R (1971) "The Minicourse - a milestone on the road to better teaching" British Journal of Educational Technology, No. 1, Vol. 2, pp 14-23

BOSLEY H E and WIGREN H E (1967) Television and Related Media in Teacher Education Multi-State Teacher Education Project, Baltimore

BRUSLING C and STUKAT K G (1972) Report on microteaching research at Gothenburg School of Education Educational Research in Sweden, 1971-72

GOLD G F (1974) "Microteaching, 1973-74: a report on some problems relating to innovation in an established course" Mimeo. Jordanhill College of Education, Glasgow

JAMIESON G H (1973) "Simulation: some implications of skills theory" Programmed Learning and Educational Technology, Vol. 10, No. 4, pp 239-247

PERLBERG A (1972) "Microteaching" International Review of Education, Vol. XVIII, No. 4, pp 548-559

ST. JOHN-BROOKS C and SPELMAN B (1973) "Microteaching" Trends in Education, No. 31, pp 14-19

WRAGG E C (1971) "The influence of feedback on teachers' performance" Educational Research, Vol. 13, No. 3, pp 218-221