Science and Mathematics Education Centre

A comparison between the effects of Keller Plan and traditional teaching methods on structure of learning outcomes among tertiary mathematics students

Mary Ruth Freislich

This thesis is presented as part of the
requirements for the award of the degree of
Doctor of Science Education
of the
Curtin University of Technology

February 1997

Abstract

The goal of the present project was to evaluate a mastery learning teaching method in mathematics for engineering undergraduates. Many mathematics teachers are very dissatisfied with the level of understanding displayed by students who pass traditional examinations. The Keller Plan requires mastery demonstrated by almost perfect performance on a sequence of tests which students repeat until they reach the high standard required.

The study compared students in the same mathematics subject in the year before a change to Keller Plan teaching, and in the year of the change. Achievement scores, defined in terms of the completeness and consistency of solutions to test problems on the whole of the syllabus, were higher for the Keller Plan group. Measures of attitudes and approaches to study, which were positively related to achievement, indicated that the Keller Plan group had stronger intrinsic motivation, and more diligent study methods. Their confidence tended to be lower than that of the traditionally taught group, but was not low in absolute, and appreciation of the greater challenge of the Keller Plan appeared to be worked out via diligence. Students felt that individual work in the Keller Plan was a better use of time than attending traditional lectures.

Acknowledgements

I should like to thank my supervisor, Associate Professor John Malone, for the help I have had from him while working on the present project. His advice has always been most enlightening, and his kindness and conscientiousness most impressive.

I should also like to thank all staff members of the Science and Mathematics Education Centre for the pleasant and stimulating contact I have had with them during my period of enrolment.

The work is dedicated to Alan, Anna and Sophie, for they know why.

Contents

List of tables

List of figures

Chapter 1. Introduction

Aims

The present work is motivated by the desire to obtain systematic evidence about alternative methods of university teaching in mathematics in the Australian context. The majority of students in tertiary mathematics classes are taught in large enrolment groups, assessed mainly by final examinations. Performance in examinations often worries markers, because it indicates that important concepts have not been understood by a very large proportion of the class. One alternative teaching method that seems applicable, and claims to produce better results, is the Keller Plan. In this method, study is individually paced, and assessment requires a very high level of mastery of material, allowing repeated test taking until mastery is achieved. A newly implemented Keller Plan subject at the University of New South Wales offered the opportunity to compare the performance and attitudes of adjacent year groups doing the same syllabus under different teaching methods.

The specific questions that the work is intended to address are quite simply stated. First, can one say that one method of teaching tends to produce higher quality observable performance than the other? And, second, is one associated with more favourable approaches to study or greater satisfaction than the other? The meaning of terms in the questions clearly requires elaboration, and such is the task of the rest of the chapter.

It should also be noted that the investigation of the main point is complicated by the existence of natural subgroups defined by gender and citizenship (a substantial minority of overseas students, mainly from South East Asia). The aims of the study must therefore include an investigation of whether such subgroups are differentially affected by the methods of teaching under consideration.

Background

Tertiary mathematics teachers tend to show incomprehension and suffering when they are faced with evidence of what their students have forgotten, failed to understand or been unable to carry through consistently when solving problems. They tend to blame students for lack of interest or laziness, to deplore the ability of their institution's intake, and to blame previous teachers for letting people pass prerequisite subjects with too low a level of performance. How much of this can one explain away as part of the ancient game of generational and intellectual class abuse?

First, there is evidence that lack of understanding is common to a wide variety of intakes, and is compatible with previous satisfactory performance in traditional examinations. Dahlgren's (1984) well known study of economics students is a good illustration. Given the question "Why does a bun cost one kronor?", university students answered in terms of intrinsic value of materials and labour, without ever mentioning the idea of supply and demand, which they produced easily in a traditional examination setting. (p. 30). Replications found quite advanced physics students using pre-Newtonian explanations, again despite good examination performance in a course on Newtonian mechanics (Dahlgren, 1984, p. 32).

There is also well established evidence that students' level of understanding depends on their conception of the task (Marton & Saljo 1976b), and their intention of global understanding in their approach to it (the deep approach as opposed to the surface approach) (Marton & Saljo, 1976a) What is more, Svensson's (1977) study indicated that diligent study habits were much easier to form among students whose habitual approach was deep, and that diligence clearly related to performance. The limitation of this mainly Swedish work is in its small and narrowly based samples. But British and Australian work on larger and more wide ranging samples, using quantitative methods (for example Ramsden & Entwistle, 1981; Biggs, 1982), found that approaches to study could be quite adequately identified by questionnaire responses, and that depth of approach could be thus detected, and that depth was related to performance. In particular, intrinsic motivation and active study methods related well to achievement, and some students seemed to be able to turn anxiety to good account. In mathematics, a high level of active work, mainly involving problem solving, is one's main evidence for the presence of diligence and interest. What is more, without some active work, intrinsic rewards are not accessible, because interest and satisfaction tend to come out of the activity of problem solving. That is, the intrinsic motivation must contain a component of what the now classic Fennema-Sherman studies of attitudes towards mathematics (Fennema & Sherman, 1977, 1978) called "effectance motivation," that is, accepting the challenge of problems and getting satisfaction from handling them. This implies also that the confidence necessary to accept challenge is both required and fostered by active study. Extended work by Fennema and Sherman led Fennema (1985) to conclude that confidence and a feeling of control over one's work were vital factors in learning to learn mathematics, at least at secondary level, where the full range of confidence levels is present among students.

It seems to follow that one cannot blame one's intake, because failure to understand appears to occur regularly at all levels. It also seems clear that most current forms of assessment are letting students pass without real understanding, so that it may not be policies about students' progress in a particular institution that need changing, but instead, methods of assessment. The conventional wisdom about idleness and lack of interest among students appears to be supported by research; that is, it is indeed both frequent and important.

In mathematics, there are particular features of the subject that make forgetting and lack of understanding fatal. The subject consists of a lattice of concepts in which each depends on a set of others, and the loss of one node of the structure often implies inability to proceed. On the other hand, the high structure of the subject, and its accessibility to reasoning mean that reconstruction of missing steps may be possible, if understanding of the general situation is good enough. What is more, rote learning in mathematics is almost impossible, because it is so difficult even to recognise a fact without some understanding, that the memory load soon becomes intolerable.

One may therefore conclude that there is great promise in any teaching method that helps students to take action for themselves, thereby opening the way to the intrinsic rewards that in turn make taking action easier, and to the level of understanding that makes reconstruction of forgotten material comparatively simple.

It is traditional in Australia to teach tertiary mathematics to large lecture groups, backed up by tutorials in groups of about 20 to 25 students, which is not, in absolute, actually very small. Material is paced by the lectures, and tutorials follow the lecture presentation. Keeping up to date is left to the students, and, because of the logical dependence of new material on old, once a student falls behind, the usefulness of classes tends to drop sharply.

It follows that alternative teaching methods that promote mastery of a topic before the student goes on to the next are well worth investigating, for the multiple reasons given above. It is clear that individual differences in ability, good organisation, learning history and present circumstances imply that the mastery requirement imposes the accompanying requirement of individual rates of progress. Hence a way of working without timetabled lectures is in turn required. One of the most a priori appealing schemes is Keller's (1968) personalised system of instruction (PSI) in which students study specially designed materials individually, and are required to show mastery of each topic before going on to the next, with mastery assessed by perfect or near perfect performance in a test that students take when they feel ready for it. The test is assessed immediately (reinforcement a la Skinner, see Keller, 1968), and, if the required level is not reached, the student must return to studying the material and take another test on it later. Grades, in the original formulation, depend on how much of the material a student can master to the required level. Keller first applied his plan in psychology classes, and made the units of material short, in accordance with the origins of the plan in Skinnerian learning theory. Hence the plan should foster confidence and continuing effort. In a tertiary mathematics framework, structure requires larger units, and the mastery requirement is far more demanding than traditional assessment, so that the effects on positive confidence are a priori more open to discussion. One must therefore look to research for guidance.

The first implementations of the Keller Plan date from the early 1960s, and the most influential paper in the field is that published by Keller himself in 1968, which was mentioned above. There has therefore been plenty of time for research, and so one must justify conducting any further research on it in any case. The first thing to note is that general reviews of research at tertiary level tend to assess PSI positively, in contrast to assessments of other forms of individualisation (for example, Kulik, Kulik and Cohen, 1979; Pearson, 1983). Students' ratings tend to be favourable in the great majority of cases, and achievement comparisons indicate that PSI tends to do as well as, or better than, the traditional lecture method.

There are, however, some difficulties about study of PSI plans in general, notably about the validity of using grades for comparisons, and about self selection, drop out rates, and greater completeness of information about test styles and content (for example, Kulik, Kulik & Carmichael, 1974; Smith, 1987). Hence overall impressions are not necessarily as solidly based as they seem.

One should also note that the use of PSI has declined since the 1970s (Lloyd & Lloyd, 1986), and that research interest dropped over a similar period (Lamal, 1984). Over more recent years, shorter lists in the indexing journals 1986-96 confirm the trend towards loss of interest. Should one therefore conclude that early good results represent a Hawthorne effect, or effects of genuine partisanship? It is worth noting that the survey by Boud, Bridge and Willoughby (1975, p.19) asks for more research by investigators "less committed to the method's success", for precisely these reasons.

On the other hand, it is possible that declining use merely reflects the strength of tradition, allegiance to "the way it spozed to be" (Herndon, 1974, sic) eventually defeating innovation. A hint of this is observable in the deliberately open and illuminative study done by Friedman, Hirschi, Parlett and Taylor (1976) at MIT, where the Keller Plan was used for a few years in a large physics class, and then discontinued without systematic investigation, probably because initial organisational difficulties had given the method a bad name that stayed with the course despite the difficulties having. been successfully cleared up.

The overall research evidence is therefore not as clear as one would wish. But a first obvious step towards greater clarity is to separate out implementations in different subject areas, because both rationales and feasibility of implementations differ by subject. What is more, most implementations have been in social science teaching. For example, Thompson (1980) estimated that only about 15 percent of studies accessible to him dealt with science and engineering, the rest being in the social sciences. Further, even when reviews restrict themselves to scientific subjects, the number dealing with mathematics is low. For example, in the early review by Kulik et al. (1974) of studies of PSI in science, there are none in mathematics, and the common indexing journals over the last decade record that only about 20 percent of the studies of PSI in tertiary science deal with mathematics. Given the special status of mathematics alluded to above, as a subject in which isolated facts tend to be unrecognisable, it seems obvious one should be cautious about comparing studies of mathematics with studies of other subjects. Even comparisons with science subjects are doubtful, because, even though the status of supposed facts is uncertain in the sciences, they are more easily recognisable than in mathematics, and missing material in the sciences is less accessible to reconstruction.

It follows that the attempt to clarify the research picture without restriction of subject is misdirected if one is interested in mathematics. But if one restricts one's attention to studies of mathematics, with the cautious inclusion of statistics, cautious because of the possibility of what is known as a "cook book" course even at tertiary level, does one find any difference, or any greater clarity in the research?

Examination of the literature shows that there is some support for the presence of differences among subjects. (Details are in Chapter 2) On the whole, results for mathematics are rather more ambivalent than those for social science, but tend to be more positive than for the hard sciences. The most enthusiastic reviews do not separate studies by subject (for example, Kulik et al., 1979) while others that do differentiate or restrict tend to be either less enthusiastic (for example, Boud et al., 1975, restricted to science and engineering) or highly aware of design problems (for example, Kulik et al., 1974, restricted to science). Restriction to mathematics alone reveals a continuing, quite substantial subgroup, containing some very well designed studies, that finds either no advantage, or some disadvantage, in PSI (for example, Howarth & Smith, 1980; Thompson, 1980; Watson, 1986a, b, and Smith, 1987). From a more recent large scale meta-analysis by Kulik, Kulik and Bangert-Drowns (1990), one can extract average effect sizes that place PSI lower in science than in mathematics, with both below social sciences.

So mathematics results do look somewhat different from those in other subjects. But are they clearer? The answer has to be no, because there are about equal numbers of favourable and neutral reports, and some of the best studies, such as that of Watson (1986a) report unfavourably. It appears, therefore, that there is a case for further study, but only if one can offer some advantage over the previous research. The most difficult problems in comparing Keller Plan teaching with traditional teaching appear to be those of self selection and of the achievement criterion.

Self selection is an obvious design problem, overcome only if groups are randomly assigned to teaching methods by the investigators. Parallel specially set up groups, however, also present problems, because knowledge of the existence of the other group may influence students. A partial solution is afforded by comparing classes just before, and just after, a point of change between teaching methods for whole year groups, and such a pair of groups was available to the present study.

The point about the achievement criterion is particularly intractable, because grades are quite obviously unsatisfactory, and a common final examination is either a betrayal of the Keller Plan group or something they need not take seriously. What is more, even in studies using a delayed examination, not part of the assessment for either group, like that of Watson (1986a), administrative difficulties can force the examination style towards the less satisfactory. Watson had to use a multiple choice test, which, in mathematics, is almost always assessed by teachers as inferior to one requiring full answers. What is more, only half the students sat the test under supervision, and, although Watson does not explicitly comment on this in the paper, the numbers involved in the post-test were considerably lower than initial sample sizes, which may or may not imply bias. The most desirable advantage in a project should therefore be in an achievement test that depends on close scrutiny of the full detail of students' work under similar conditions. Such work should be complete enough for degrees of integration and consistency, as required for mastery, to be observable. That is, one can justify further research if one makes the comparison hard and detailed.

The present project is fortunate in having access to a newly implemented Keller Plan course that grew out of very long experience in a well established Keller Plan implementation, the same one studied by Barrett and Prokhovnik (1980). The established course therefore has afforded long experience that transfers to relatively trouble free implementation of the new one. The subject is compulsory for chemical engineering students, so self selection is not a danger. Obtaining the first year of implementation means that students from the previous year were still in the university for inclusion in a control group. Therefore the crucial factor is the achievement criterion.

One should note that, if one does not use grades or a post test, there is in any case a problem about assessing achievement within a Keller Plan course, because students do different tests, even though on the same material. One therefore requires a demonstration that there are strong regularities in observable outcomes of learning as students' knowledge develops. The idea of structural hierarchies, as in the Piagetian model, is clearly relevant here, but one wants something more modest than Piaget, involving stages of learning in a particular context rather than stages of overall intellectual development.

The more recent constructivist writers, such as Cobern (1993) tend to reject the idea of linear stages of development, in favour of focussing on the actual content of individual students' thinking about a topic, presumably meaning observable indicators of such content. It does not, however, follow automatically that there is no regularity in observable solutions of problems or answers to questions: one needs evidence for that. Is there, therefore, work in the literature that encourages one to look, in students' actual work, for structural commonalities that have some justifiable ranking, without looking for universality or strict linearity of overall stages? The research strand to be examined next does offer such encouragement.

The strand emerges from the interaction between the approach of Biggs and Collis (1982), which originated in the idea of refining and operationalising regularities in study outcomes in the Piagetian tradition, and that of Marton and his colleagues at Göteborg (Marton & Säljö, 1976a, b; Svensson, 1977, Fransson, 1977), which was mentioned above in connection with students' conception of, and approach to, learning tasks. The Göteborg group focussed on the quality of learning outcomes in a particular task done by a particular sample of students. Within this limited universe, they made multiple independent studies of processes, intentions and structure of outcomes, which produced extremely coherent results. Briefly, independent classifications of approaches and responses came out with almost perfectly consistent divisions between deep (intending to understand the whole task) and surface (just accumulating unstructured detail) approaches (Marton and Säljö, 1976a, b), and between levels of integration within the actual responses produced (Svensson, 1977). That is, students who intended to understand the task as a whole tended to produce much more integrated and coherent responses to questions on the material learned, and agreement between independently assigned categories was almost perfect. The Göteborg group was very careful to claim validity of their classifications only within the confines of the particular interviews used, and the particular responses produced, in the study. Their followers, however, went further, as has been indicated above in connection with questionnaire development for assessing approaches to study. The aspect involving quality of observable outcome was also developed and extended by subsequent research.

Biggs and Collis (1982) extended the work on classification of observable outcomes of learning according to their structure, analysing students' responses to questions for elaboration, consistency and extent. Their work differed from the Swedish work in that it attempted to go beyond particularity of task and group, with the intention of providing gradations of learning outcomes valid across different tasks, subject matter and student groups. They formulated a set of classificatory rules defining the Structure of the Observed Learning Outcome (SOLO) taxonomy, and claimed that it produced equivalent structures in all subject areas. The actual examples they give are from early secondary school level. Unfortunately, in several areas, the status of the evidence underlying the examples is not made clear, and in some subjects the higher level examples seem rather artificial. The work on mathematics, however, is very clearly based, being taken from a sequence of extensively documented studies done by Collis (1972, 1975a, b), about the classification of levels of mathematical thinking.

The point of interest for the present work lies precisely in the extended range of the classification, with the simplifying and facilitating feature that it is mathematical work within a restricted range that is important. This means that one is working from the initial definition for which the quality of the evidence is clearest, and that it is not necessary to become involved in arguments about total generality of classification. Initially, there was some confusion in the literature dealing with the taxonomy. Biggs and Collis (1982, pp. 24-25) appear at first to have intended an operationalisation of a Piagetian hierarchy, but in work that uses the taxonomy as a framework in mathematics education, such as that of Pegg and Davey (1989) and Pegg (1991), the SOLO levels tend to be taken as refinements within Piagetian levels, and considerable emphasis is placed on performance in separate areas, as opposed to the overall premiss in Piaget's work.

The criteria defining the SOLO levels are those of completeness, consistency and integration of solutions to problems. It is clear that such criteria have much in common with the Goteborg classification of work as showing deep processing and holistic learning. The very top level proposed is, in mathematics at least, out of range for undergraduate work, being very much at the flexible and creative level required for research. The first four, on the other hand, suggest a very reasonable way of ranking solutions to a problem. They are defined as follows.

1. Prestructural: essentially no valid response.

2. Unistructural: one aspect of the problem correctly identified, but without any diversity of aspects recognised, and hence no development in the solution, and, a fortiori, no question of consistency arising.

3. Multistructural: multiple information recognised as relevant, but not all relationships between elements clear, so that inconsistency occurs.

4. Relational: multiple information brought in and used consistently, thus with relationships relevant within the task recognised and used.

The work of Pegg and Davey (1989) found that SOLO categories gave a good representation of students' actual work at lower secondary level in geometry. The work is of particular interest, because it compared the usefulness of the SOLO levels with the best known alternative in geometry, the five levels of functioning defined by Van Hiele (1986), to some degree similar to SOLO, but found by Pegg and Davey to be slightly less useful, because less finely graded at lower levels.

Chick, Watson and Collis (1988) found that SOLO levels could easily be adapted to an analysis of mathematical work produced by first year trainee teachers. They emphasise here that tasks may limit the complexity of required solutions, so that task analysis is of great importance. One may note that task analysis is a prerequisite for comparisons between work on different tasks, as required for the present project. A full discussion of relevant research is in Chapter 2.

Pilot work for the present project found no difficulty in applying SOLO categories to the written mathematics work of university students in engineering. It was therefore decided to use such methods to establish achievement scores for use in the project, and thus to provide the severe test of the Keller Plan studied, as compared with traditionally taught groups in the same sort of area. A full explanation is in Chapter 3, and supporting evidence is in Appendix 1.

Methods

Given the above discussion of criteria for analysing students' written performance, it seemed possible not to waste information available for total enrolment groups. This, in turn, implied that it would be wasteful not to include relatively large samples in the study of attitudinal factors. The larger part of the present work therefore used quantitative assessment of data from written materials. That is, both achievement scores from assessment material and attitude scores from questionnaires were obtained for large groups, and subjected to statistical analyses. Because such methods may always allow nuances and unforeseen salient issues to be missed, a small amount of supplementary information from open ended interviews was also collected. Methods and samples are fully described in Chapter 3.

Implications of the work

The advantages claimed for the Keller Plan are such that it seems highly valuable to obtain further systematic evidence on their justifiability in a given context. The context of the present work is of particular importance, for two reasons. First, the research pattern for tertiary mathematics classes does require clarification. Second, it is of special importance to test claims using a large class of engineering students, who may be considered as the most obviously intractable material. If higher quality learning in the plan can be established, there are stronger arguments for change of organisation. Results in the opposite direction, on the other hand, would confirm Watson's (1986) more pessimistic assessment of the plan, and imply some consensus in this direction, at least in the context of Australian tertiary mathematics teaching. Because implementing the Keller Plan is economically feasible in this context, such discrimination has clear practical consequences. Similar remarks apply to attitudinal differences, subject to an argument in context about relativities between them and achievement results.

The claims, their research backing, and their applicability to tertiary mathematics teaching, are examined in detail in Chapter 2.

Chapter 2. Survey of literature

Introduction

In what follows, the scope of the review of literature is focussed on studies of Keller Plan implementations only, not other forms of individualised learning programmes. It is also, when studies of different teaching methods are in question, largely restricted to the teaching of mathematics. It is argued that this is necessary because there are both intrinsic arguments and research evidence to support the claim that mathematics learning has different requirements and different implications, compared with learning in other subject areas. This means that results on processes of learning that cover a wide range of subjects may have to be adapted to improve their relevance to mathematics. It is also clear that only studies of Keller Plans involving tertiary students are really directly relevant, although more general work on mathematics learning, involving students of all ages, may have contributions to make.

The review therefore considers in detail first the available studies that are specifically concerned with Keller implementations, in mathematics, at tertiary level. Since the results for mathematics are not very clear, it is argued that a somewhat different approach to evaluating achievement is required.

The approach that suggests itself originates in work done in Sweden, and, later, in Britain and Australia, on students' approaches to learning and their relationship with the structure of observable outcomes of learning. It is argued that the work on approaches needs modification to be useful in connection with mathematics learning, and can benefit from results for secondary students, such as the findings of the now classic Fennema-Sherman studies.

The Australian work on structures of outcome that is most relevant, is the construction by Biggs and Collis (1982) of the SOLO taxonomy, whose applicability to earlier mathematics performance is particularly well researched, and whose basic principles are not difficult to adapt to more advanced material.

The work described immediately above gives the basis for the choice of attitudinal factors to be investigated, and for the construction of achievement scores to be compared.

There are reasons to believe that the effects of teaching methods may be different for different subgroups of the student body. In particular, cultural differences, levels of selection, and gender, are relevant to the present study. Culture and selection are relevant, because Australian universities, and particularly engineering faculties, contain a substantial minority of overseas students, mainly from South East Asia, who are clearly from a different background.

This group also tends to have a rather different pre-university experience from that of most Australian students. They are also rather more highly selected than the local group with respect to previous performance. It is argued below that separate consideration of the group is essential, but that there are not good guidelines in available research for prediction of patterns of expected differences.

Gender is relevant because, where there are groups of women to be considered, apart from the general interest arising from a naturally occurring dichotomy, there are quite adequate grounds in existing research for regarding gender as an important issue in mathematics learning. Because the number of female engineering students is steadily increasing, there is usually a sufficiently large group of women to be considered, and, among chemical engineering students, the principal group being studied, the female subgroup is not of negligible size. It is not clear, however, which set of findings in the literature is applicable because of the specialisation of the group, and this is briefly discussed.

Reasons for restriction of scope

Keller first developed his teaching plan, also called the Personalised System of Instruction (PSI), in the early 1960s. Judging by citations in papers about implementations, it became widely known and used mainly as a result of his 1968 paper "Goodbye teacher...". There has therefore clearly been plenty of time for research to evaluate the plan.

Research interest, which one can assess by looking at the indexing journals, appears to have been high in the 1970s, but subject to a later decline, although it does not fade out completely. Lamal (1984), looking at indices of conference proceedings for reports on studies of the Keller Plan, found a similar pattern in the incidence of such reports. Lloyd and Lloyd (1986) found the same pattern in the number of reported implementations of the plan. A small survey by Karp (1983) found that, among university teachers who had used the Keller Plan, only just over a quarter were still using it.

Without more specific information, one might conclude, from the above information, that experience had been unfavourable, leading to loss of interest. But the findings of the research point in the opposite direction. That is, if one looks at the conclusions reached in many reviews, in particular the continuing sequence produced by the Kuliks and their colleagues (Kulik, Kulik & Carmichael, 1974; Kulik, Kulik & Smith, 1976; Kulik, Kulik & Cohen, 1979; Kulik, Kulik & Bangert-Drowns, 1990) one would be led to the opposite conclusion, that is, that research has declined because there is so much evidence in favour of the Keller Plan that further work would be redundant.

There is therefore continuing conflict between enthusiastic research reviews and the decline in the number of implementations. There is also some evidence that suggests that early enthusiasm could have led to a Hawthorne effect, with the decline in interest accompanying a decline in effect both on teachers and on students, due to the inevitable decline in the novelty of the plan. For example one review (Boud, Bridge & Willoughby, 1975) saw much of the research as somewhat contaminated by the research workers' already existing strong commitment to the success of the Keller Plan. What is more, there is considerable evidence that schemes involving individualised instruction tend to be subject to diminishing returns over time. For example, a review by Horak (1981), dealing with several different methods of individualising school mathematics teaching, found a strong tendency for effects to diminish over time, as one might expect if initial enthusiasm among implementers had been a factor. Similar results were found by Romizowski (1979), reporting on individualisation plans that included some at tertiary levels:

> . . . in large institutional applications . . . results deteriorate with time (p. 148)

Romiszowski proposes that the pattern of early success and later decline is due to an initially greater effort by teachers, which wears off over time. It is worth noting that students' attitudes to Keller Plan implementations have sometimes been found to be highly favourable initially, but showing declining favourability over time (see, for example, Roth, 1973).

But one should be aware of influences towards the abandonment of Keller Plan implementations that are the reverse of a Hawthorne effect, a condemnation stemming from the mystique of tradition, that eliminates innovations because they conflict with the supposedly correct way of doing things. The case study of a Keller Plan implementation in physics at MIT, presented by Friedman, Hirschi, Parlett & Taylor (1976) fits such a pattern. The implementation involved, once written up with overwhelming enthusiasm by Green (1971), appears to have been abandoned because of the continuing influence of initial bad publicity, whose causes had in fact already been remedied, rather than as a result of a carefully considered assessment. It is also possible that unsuitable instruments may be used to assess implementations, as found by Hassett (1978), where different levels of favourability in attitudes were found using different instruments.

One should also be aware that individualisation in the Keller Plan is not present purely for the sake of allowing individualised study, but instead, in the service of the mastery requirement, which is intrinsic to the quality of learning, so that one cannot extrapolate too freely from studies of varying types of individualisation.

There is therefore a considerable amount of uncertainty about the conclusions one can draw from reviews of research, in the context of varying types of individualisation. In the attempt to clear up confusion, one has to restrict scope and look at detail.

A first restriction comes from evidence that confirms the suggestion made above, that Keller Plan implementations must be considered separately from other forms of individualised instruction. Two studies that distinguish the Keller Plan from other forms of individualisation, and make comparisons between effects (Kulik et al., 1979; Pearson, 1983) in fact find that the Keller Plan tends to produce better results. It therefore seems essential to be quite strict about the admissibility of studies, by restricting to those that deal with teaching methods that do not break the essential form of the Keller Plan.

Since implementations do tend to vary in detail, one must decide what is essential. There is a body of research that examines the contribution of separate elements of the plan. In the review by Kulik, Jaska and Kulik (1978), there are studies referred to, such as that of Calhoun (1973), that conclude that each element of Keller's original definition makes a separate contribution to the success of the plan. On the other hand, there is evidence collected by Powell (1980) that indicates that the most important feature is the mastery learning requirement, that one need not be too strict about the length of units, that a certain amount of pressure in pacing is acceptable, and that student proctors are not vital. Semb (1981) also argues that mastery learning is the core of the plan, with the other features serving to facilitate it. The argument that the length of units is not uniformly determined is supported by research done by Spencer and Semb (1978), who found that students who were already coping well under PSI could use longer units without their measured achievement suffering. In later statements, Keller himself states that there is no formula for unit length, and that the nature of the material dominates (Keller & Sherman, 1974). There is also evidence from the review by Imrie, Blithe and Johnston (1980) that some forcing of pace is so common as to be regarded as normal. The conclusion drawn here is that one can be flexible about features other than mastery learning and its concomitant multiplicity of test opportunities with rapid feedback to the student. This means that the implementation studied in the present work is considered not to fracture the essential definition, and that research is to be considered relevant if it deals with plans that retain the essential mastery requirement and testing programme, allowing some flexibility about other features. Thus in the present work, since it is tertiary mathematics learning, above the introductory level, that is being studied, the feature prescribing the use of student proctors to assess tests is just too hard to be realised, because the level of material and speed and accuracy in marking required are at a level beyond that of any available population of potential student proctors, but the course is still considered admissible as a Keller Plan implementation. It is worth noting that many implementations include a similar decision.

The questions of level and subject area have still to be resolved. It seems reasonable to restrict consideration of research that deals directly with Keller Plan implementations to studies at tertiary level, because the present work is on tertiary studies, and inclusion of other levels brings in unnecessarily confusing degrees of freedom. On the other hand, in considering design questions for the present study, it is obviously legitimate to consider more general work, for example, general results about student learning at tertiary level, or about mathematics learning at various levels in a range of populations.

It is also important to restrict consideration of studies of Keller Plan implementations to those dealing with the teaching of mathematics. There are two principal sets of reasons. The first arises from the same characteristics of the subject that impose longer units. Mathematics is close to unique in the cohesion of whole topics and in its inaccessibility to rote learning. This means that, especially at tertiary level, logical dependence of arguments is so strong that indecomposable teachable blocks tend to be relatively large, and very low level atomistic learning is almost impossible. That is, it is almost impossible even to recognise a fact in mathematics without some understanding, and understanding ranges beyond the individual steps of argument. This means that Keller Plans in mathematics cannot be as closely related to the origins of the plan in programmed learning as they may be in other subjects, even, in certain components, quite high level science subjects. What is more, the convergence of mathematical arguments is very strong, and the level of argument, at least in undergraduate subjects, is strongly determined by the problem, so that mastery at a given level is particularly easily defined. There are therefore strong intrinsic reasons for regarding Keller plans in mathematics as somewhat specialised, so that comparisons with studies of other subjects are inappropriate.

The second type of consideration is imposed by aspects of the research literature. First, there is some evidence that supports the argument that studies of mathematics learning are likely to differ from studies of other subjects. In the recent large meta-analysis done by Kulik et al. (1990), the main subject areas are mathematics and statistics, the physical and biological sciences, and the social sciences. The analysis is based on average effect sizes, and there is enough information to permit separate calculations by the reader of average effect sizes in the three areas mentioned. The average effect size for PSI in social science is .68, and that for physical and biological sciences .26, while that for mathematics is in between at .44. Such differences are consonant with the argument that separation should at least be considered.

The argument for it is further strengthened when one looks at the distribution of studies over subjects. The original Keller Plan was designed for a psychology subject, and review material tends to show a much larger number of studies of implementations in the social sciences. For example, in the set reviewed by Van der Klauw and Plomp (1974), there are 250 studies of psychology courses as opposed to 27 in mathematics. This means that, in a general review, studies of mathematics would be likely to be swamped by studies of subjects whose fit with the Keller rationale is different enough to mislead an enquiry into mathematics learning. Similar remarks apply to studies of learning in the physical and biological sciences, which, though less plentiful, are still frequent enough to weight conclusions if combined with studies of mathematics. The detailed review of studies of Keller Plan implementations that follows is therefore restricted to work on mathematics. This means that results of general reviews must be discounted unless their information is sufficient to allow mathematics to be separated out.

Studies of Keller Plan mathematics courses

Achievement comparisons

A first point of interest comes from an examination of sources used in reviews. To be specific, one can look at the thirteen studies of mathematics and statistics used in the meta-analysis carried out by Kulik et al. (1990). These are tabulated, with their sources, in Table 2.1.

Simple counting reveals that less than a third of the reports come from refereed journals. The problem with the others is that different reporting conventions or lack of space mean that it is hard to assess the quality of work reported outside journals. In particular, Dissertation Abstracts seldom gives space for complete reports, and conference papers tend to be both lacking in detail and hard to access. Kulik et al. do claim to have selected their material to avoid methodological flaws, but it is hard to see how their sources would permit soundly based selection. It follows that the findings of the meta-analysis are somewhat weakened because of the presence of studies with restricted sources.

It follows that, in the present review of material, it will be considered very important to separate material with sources complete enough to allow assessment of quality from materials whose sources are too restricted to permit assessment, so that conclusions may be appropriately weighted.

Table 2.1 Sources of studies of mathematics reviewed by Kulik et al. (1990)

Study	Date	Source
Condo	1974	Conference paper
Kulik and Kulik	1976	Book chapter
Locksley	1977	Dissertation Abstracts
Lu	1976	Book chapter
Malec	1975	Journal
Nord	1975	Dissertation Abstracts
Pascarella	1977	Journal
Peluso and Baranchik	1977	Journal
Schielak	1983	Dissertation Abstracts
Smith	1976	Dissertation Abstracts
Steele	1974	Conference paper
Taylor	1977	Conference paper
Thompson	1980	Journal
White	1974	Book chapter

The results of the literature search for studies of achievement are tabulated in Tables 2.2 and 2.3, and the information contained in them is discussed below.

Table 2.2 contains material whose source is considered sufficiently detailed to allow assessment of the quality of the work. It contains, first, studies found in a journal search made for the current project, subject to the elimination of studies found to have totally inappropriate methods. For example, the study by Studer (1976) was eliminated because it made no tests of the significance of results. On the other hand, that of Rae (1993), which also gave no tests, did display sequences of mean examination marks, and hence was not eliminated because of the absence of testing, even though the plan used also differed in some essential features from a Keller Plan implementation. In addition to journal articles, the table includes book chapters from Kulik's list, subject to availability and clarity. A third set of studies admitted consists of those documents from Resources In Education that are complete enough to reveal the design of the study. Any reservations about study design are noted in the table.

Table 2.2 Studies of achievement in Keller Plans compared with traditional teaching: full reports available

Study	Control	Comparison	Results
1. Harris & Liguori (1974)	Pretest.	Final examination, scored with one point per question, which probably means multiple choice or very small units.	Not significant.
2.Klopfenstein (1977)	Previous achievement and test of mathematical aptitude. The Keller group had higher scores, but significance is not given	Final examination, described as having 20 items, which probably means multiple choice or very small units	Not significant
3.Malec (1975)	Not clear. Cohort study.	Final examination, type not clear.	Favoured Keller, on one-tailed test
4.Martinez & Martinez (1988).	Students assigned to groups, but the Keller Plan groups were taught by volunteers.	Final examination, type not clear.	Not significant
5. Pascarella (1977).	Random selection of sample, but within self-selected groups.	Final examination, full answers.	Favoured Keller
6. Pascarella (1978): extension of 1977 study.	Pretest, no significant differences.	Final examination, full answers.	Favoured Keller
7. Peluso & Baranchik (1977), first study.	Pretest, no significant differences.	Final examination, type not clear.	Favoured Keller
8. Peluso & Baranchik (1977) second study.	Pretest, Keller group lower. Note Keller group had higher withdrawal rate.	Final examination, type not clear.	Mixed
9. Rae (1993)	Previous years' examination results	Final examination, full answers	Keller mean high compared with traditional norm.
10. Smith (1987).	GPA, IQ, no significant differences	Final examination, type not clear	Not significant
11. Struik & Flexner (1977).	GPA, mathematical ability tests used as covariates. Note Keller group had higher withdrawal rate.	Final examination, full answers, but with groups separately marked.	Favoured Keller
12. Thompson (1980).	Random assignment of students and teachers.	Final examination, standardised multiple choice test plus internally set test requiring full answer	Not significant

Study	Control	Comparison	Results
13. Van der Klauw & Plomp (1974)	Not clear. Claimed groups were "of equivalent size and characteristics".	Final examination, form not clear, but described as "usual", which probably means full answers.	Not significant.
14. Watson (1986).	Previous achievement and pretest as covariates.	Multiple choice post-test, to assess retention.	Favoured non-Keller

Table 2.3 contains results from conference papers and very short documents, together with reports from <u>Dissertation Abstracts</u>. This list is given lower importance than that of Table 2.2, precisely because of the restrictions in the sources of information about the studies listed.

Table 2.3 <u>Studies of achievement in Keller Plans compared with traditional teaching: restricted reports.</u>

Study	Control	Comparison	Results
1. Collard (1989).	Pretest.	Final examination, type not clear, gain scores.	Not significant
2. Edward (1976).	Pretest, standardised, so probably multiple choice. Used as covariate.	Post test, similar instrument	Favoured Keller
3. Hassett & Thompson (1978).	Claimed good, no other information.	Final examination, type not clear.	Favoured Keller
4. Kontogianes (1974).	Not clear.	Final examination, type not clear, gain scores.	Favoured Keller
5. Locksley (1977).	Not stated.	Final examination, type not clear	Not significant
6. Nord (1975).	Background variables only.	Final examination, type not clear	Claims Keller Plan is a "viable alternative"
7. Rainey (1980).	Not clear.	Final examination, type not clear.	Not significant
8. Schielak (1983).	Reasoning test, arithmetic test, past course taking and random assignment to groups.	Final examination, 50 items, so probably multiple choice. Transfer test	Mixed
9. Sheehy (1990).	Random assignment within timetable constraints.	Final examination, type not clear	Favoured Keller
10. Urban (1972).	Random assignment.	Final examination, type not clear	Favoured Keller
11. Urbatsch (1980).	Matching, method not clear.	Final examination, type not clear	Not significant
12. Vatanavigkit (1985).	Pretest.	Final examination, multiple choice.	Favoured Keller

If one just counts different reported results, a conclusion in line with the enthusiastic reviews, but somewhat more restrained, seems to be required. That is, the Keller Plan very seldom does worse than traditional teaching methods in tertiary mathematics, and, almost half the time, better, at least as measured by the examinations used by the studies. But simple counting assumes equal weight, so one has to look more closely at the studies.

It should be noted, first, that the restricted reports do indeed cause problems, because quite often essential information is not given. Studies that do not clearly identify the actual teaching methods they examine, such as that of Kerrigan (1976), were omitted. But in those that survive this first test, important questions such as control are often not clearly covered. In the list of restricted reports, control is in doubt in the cases of Kontogianes (1974), Locksley (1977), Rainey (1980) and Urbatsch (1980), which means that the weighting given to these studies must be reduced. In some cases, such as that of Nord (1975), the type of control is clearly described, and shows as clearly inadequate, since only background variables were checked. Sheehy's method of assignment of students to classes cannot be established as random on the basis of the information given, because it was subject to timetable constraints.

One of the studies, that of Rae (1993), which does give a full report did not have an explicit control group, but is admitted, though somewhat reluctantly, because it appeals to a norm established by long practice.

But in general, problems involving control are also present in the rest of the set of studies that do sufficiently specify the type of control, which also makes results less clear than one might conclude from the table. Details of such studies in the set that gave full reports are as follows. Harris and Liguori (1974) found that their Keller Plan group had lower scores on the pretest, and so results for this group are open to reclassification upwards. A similar alteration of results might be needed in the study by Struik and Flexner (1977), where pretest results showed a difference in the opposite direction. One should note that the studies that are also in the set reviewed by Kulik et al. (1990) range from totally unclear about control, through flawed to good, without this receiving any specific comment in the review paper.

Among the studies reporting mixed results, one must lower the weighting given to that of Martinez and Martinez (1988), because of the design defect resulting from allowing volunteer teachers for the Keller Plan group, which was of verifiable importance because the study found strongly detectable teacher effects.

Both sets of reports involve a further difficulty involving supply of information, because they are often far from clear about the type of test used for achievement comparisons, and this also makes evaluation of their quality more difficult. It is clear that, in studies of the Keller Plan, achievement comparisons must involve scrutiny of the students' work, rather than just final grades, because grades are likely to be given on different bases in the different types of subject organisation. The studies here all satisfy that condition, but, in 16 cases out of 26 (both lists together) the type of test used for comparison is not clearly described. Ideally, one would want access to a complete test as the raw material for assessing the value of a study, which is obviously a demand unlikely to be met. But even without full details, there is a clear point to be made about types of test. That is, in mathematics, it is very important to know whether a test had multiple choice questions, as opposed to requiring full answers. This is because mathematicians usually regard multiple choice examinations as decidedly inferior to those asking for full answers, for the following reasons. First, it is the thinking behind the final answer that is important, so that the presence of a full argument gives a better idea of a student's level of performance. Second, the usual 20-25 percent chance of a right answer being randomly chosen tends to distort distributions. It is therefore important to be able to discriminate at least between the two types of test, so it is a real disadvantage that it is impossible to do this for nearly two thirds of the studies covered by the two lists. It is, one would also predict, particularly important when assessing Keller Plan teaching. In this connection the remarks made by Casanova, Casanova, Fernandez and Villamanan (1979, p. 67), about physics students, are particularly illuminating:

> although these [Keller Plan] students had a better grasp of how to solve problems, they were worse in actual calculations and problem solving than their counterparts who had attended regular problem sessions.

It is worth noting, in this connection that some of the best studies encountered used multiple choice tests, presumably for administrative reasons. For example, among those fully reported, the Australian work of Watson (1986) had to use a multiple choice post test for a discrete modelling subject. The area is often found hard by students, and the availability of partial answers and the visibility of clues would usually be thought to afford better material than is possible to obtain with multiple choice questions, as a basis for assessing the quality of students' learning. This is one specific consideration that favoured the inclusion of Rae's (1993) study in the review, despite its lack of an explicit comparison, because its achievement measure used full answers to questions in a similarly difficult field. Among the restricted reports, that of Vatanavigkit (1985) describes multiple choice examination of a calculus course, to which similar remarks apply. Thus even among the best designed studies, a very commonly used type of test is likely not be the best instrument for the task undertaken.

What is more, in the best designed work, there is an overall lack of clarity in the total set of results, because all positions are to be found, from negative (Watson, 1986) through neutral (Thompson, 1980) to intermediate (Smith, 1987) and positive (Pascarella, 1977, 1978). Among the restricted reports, those that have assessably fair to good control are more often positive (Edward, 1976; Schielak, 1983; Vatanavigkit, 1985) than neutral (Collard, 1989).

Results involving achievement are therefore of a value somewhat restricted by doubts about control and instruments. The unqualified enthusiasm of certain general reviews can therefore not be applied to studies of tertiary mathematics learning. One can, however, confidently say that results clearly unfavourable to Keller Plan implementations are quite rare, with Watson's findings isolated, even though the good design of Watson's study gives it considerable importance. Here it is worth noting that clearly negative results are also rare in tertiary level implementations in Science-based subjects in general, although there are some relatively isolated unfavourable cases like the studies of engineering students by Roberson and Crowe (1975) and Friedman, Kaplan and Cheatham (1979) .

But given the criticisms of method and reporting described above, and given the mixed finding of studies in tertiary mathematics, there does seem to be room for further work. It would, however, not clarify the picture much if it could not overcome some of the disadvantages that seem most important in the previous work.

One would therefore require samples whose previous achievement in similar tasks was known, and best of all, whose previous preparation was similar. Here one may note that these conditions are satisfied by the target group of the present work, consisting of engineering students at the University of New South Wales, who come out of a common required first year mathematics subject.

One would also hope to do better than multiple choice scores for the final achievement criterion, because that is a clear limitation on the value of work at all rating levels with respect to the rest of the design. The arguments above indicate that the tasks used for the construction of an improved score should require full answers, and that scoring should be based on close scrutiny of the detail of students' work. And if possible, stringent criteria of quality evaluation are required.

The background in the literature to the planning of a test intended to satisfy, at least partially, the above requirements, is discussed almost immediately below. But before moving on to this, one should point out what sort of inference one could draw from the results of a test considered reasonably adequate. A positive result would imply that the use of full answers would help to overcome doubts due to Watson's results, and would allow one to accept to a greater extent the favourable claims in the area. A negative result, on the other hand, would confirm a sceptical approach to such claims in the case of tertiary mathematics. But a neutral result would be highly unsatisfying, unless one had evidence other than the achievement results that could be used to resolve the question. The obvious alternative area of investigation is students' attitudes, and such an investigation is an essential part of the work on which the proposed test design is based. Attitudes to the subject have often been considered in connection with Keller implementations, and it is obviously appropriate to consider these, and results will be discussed later in this chapter. But it is the area of students' approaches to their studies, and the connection of approach with outcome that is of greatest interest, and this is what is discussed immediately below.

Evaluating the structure of learning outcomes

The idea of evaluating the quality of students' learning via a classification of the structure of the work they produce has been developed in two strands of research. These are concerned with how students learn, and with the attitudes and approaches to study that have come, as a result, to be considered the matrix of high quality learning. The work originated in the early to mid 1970s, came to unity and wide coverage in the late 1970s and early 1980s, and has since merged with work specific to mathematics learning in the generation of recent research on students' mathematical development.

The crucial section of the work came from a group at the University of Göteborg. Marton and Säljö (1976a) analysed students' answers to questions about the content of a brief article they had been given to read. They found that students tended to polarise into two groups, one identified as showing deep processing of the material, and the other identified as showing surface processing. Deep processing was evident in answers that showed a grasp of the overall structure of the argument, and had succeeded in integrating the details they mentioned into this structure. Surface processing was identified when the answer gave only a set of isolated factual details from the article, with no sign of integration of what had been learned. The dichotomy, it was argued, illustrated differences in the quality of learning, often not identified by achievement tasks whose emphasis was on quantity. The report of the research gave copious illustrations from students' answers, which made it very clear what the criteria for the classification were, and why the dichotomy was justified.

The work has much in common with British work done at about the same time, on learning styles. In a very well known paper, Pask (1976) distinguished between holistic and atomistic learning styles, on the basis of how far global understanding was attempted. For the purposes of the present project, however, an account of the Swedish work is preferred, because it is much clearer, and contributed more directly to the related work on approaches to study among much larger samples of students, which became a very important development in work on student learning.

The Göteborg group also did work important to the establishment of the connection between outcomes and attitudes. Marton and Säljö (1976b) interviewed the sample they used in their first study, about how they approached the learning task it involved, and found a clear relationship between approach and outcome. Deep processing required a clear intention to look for overall understanding, an attempt to grasp the writer's purpose and how it was worked out in the detail of the argument. The students identified as having used surface processing showed no such intention. Levels of processing were therefore linked to different conceptions of the task.

Connections with independently measured attitudinal factors were also studied by the group. The work of Fransson (1977) linked intrinsic motivation and low anxiety to higher quality of learning. Svensson (1977) found that students who habitually used deep level processing tended to find it easier to form active and diligent study habits, and that these were positively linked to achievement in the expected way. The point about forming active habits appears to rest on the thesis that surface level processing is like learning a list, and hence makes study very much more boring than it is for students in pursuit of overall structure and meaning. The relevance of these considerations to mathematics learning is obvious, because of the extreme inaccessibility of mathematics to rote learning.

The limitations of the Göteborg work are in the small and restricted samples used. Restriction here means restriction of samples to particular types of student, because all sample members were studying social sciences, and the overwhelming majority were female. For the style of work, which used close individual contact with the students, numbers are not small, but wider reference and educational usefulness require further development. The disadvantages due to restriction were, however, rapidly overcome in the research activity that grew out of the Swedish work, and adapted its insights to methods that could be, and were, used with economy on very wide ranging samples.

Two sets of studies are relevant here. The first, starting with Australian work by Biggs and his colleagues, most notably that reported by Biggs and Collis (1982), takes up and extends the idea of quality of learning outcome. They use as quality criterion the degree of integration and the extent of the links between different parts of the required material that could be found in students' answers to questions and solutions to problems.

The second relevant set of studies consists of two subgroups. One consists mainly of British work, centred on that of Entwistle and his colleagues. The development of the work is to be found in the reports of Entwistle and Wilson (1977), Entwistle, Hanley and Hounsell (1979) and Ramsden and Entwistle (1981). Conclusions are collected and interrelated in the work of Entwistle and Ramsden (1983), and Marton, Hounsell and Entwistle (1984).

The other subgroup contains Australian work, again done initially by Biggs and his colleagues (Biggs, 1976; 1979a, b; 1982). This strand builds on the studies that indicated that attitudinal factors could be regarded as the matrix of quality of learning outcome.

The two subgroups of studies reached a level of agreement that was quite startling in view of their independent origins. The work has been developed further, for example by Entwistle and Waterston (1988), and is reviewed by Entwistle (1992), but it is the original formulation that is of interest to the present work

The relevant material from the two sets of studies is outlined immediately below.

Classifying learning outcomes

Logical priority goes to direct examination of observable evidence of quality of learning . Biggs and Collis (1982) developed their classification of students' work on a variety of educational tasks, with reference to complexity, extent of coverage, and consistency of the observable structure of responses. The resulting classification was hierarchical, and analogous to the Piagetian classification of stages of development in thinking. The levels obtained by Biggs and Collis were the basis of their Structure of Observed Learning Outcomes (SOLO) Taxonomy. The relationship with Piaget's modes of thinking is not very clearly determined in Biggs' and Collis' presentation, but a rough parallelism between modes and levels of outcome seems to be part of the argument. In later work, such as that of Pegg and Davey (1989), however, SOLO levels are considered to be identifiable within modes, and to be rather more task specific than modes. For the purposes of the present work, careful resolution of that question is not important, but the latter view seems more relevant.

Biggs and Collis (1982) claim that the SOLO taxonomy is invariant across disciplines, and present illustrations, in the main areas of school study, at about upper primary to junior secondary level. Across the board, they do not make the source of the evidence clear, and some of the examples have the flavour of invention. But the mathematics work is solidly founded on research by Collis (1972; 1975a, b) into patterns of errors and misconceptions in mathematics learning. What is more, the identification of types of answer structure is easily applicable to more advanced mathematical tasks. The development of achievement classifications in the style of the SOLO levels therefore seems possible and potentially useful in a variety of contexts and at several levels.

The classification offered by Biggs and Collis (1982) is summarised in Table 2.4.

Table 2.4. Outline of SOLO levels.

Level	Definition
1. Prestructural	Essentially no valid response
2. Unistructural	One aspect of the problem correctly identified, but no diversity of aspects presented, so that questions of consistency cannot arise.
3. Multistructural	Multiple relevant information presented and used, but without considering relationships between different parts, so that inconsistency appears.
4. Relational	Multiple relevant information presented and used in a way that recognises relationships and achieves consistency within the given task.
5. Extended Abstract	Multiplicity recognised and consistency achieved over a context beyond that of the given task.

More recent research has gone further in investigating the fit of the SOLO taxonomy with students' formulation of answers to mathematical questions. Regularities in the development of geometrical concepts were also proposed by Van Hiele (1986), and the work of Pegg and Davey (1989) examined the fit of both this classification and the SOLO taxonomy with the work of junior secondary students. Results indicated that the SOLO taxonomy afforded a good classification, which was rather better than that offered by Van Hiele's system because it allowed for finer gradations at lower levels.

Chick, Watson and Collis (1988) extended the evidence that SOLO levels give an adequate classification of students' mathematical work by fitting them to problem solutions given by first year trainee teachers.

Preliminary work for the present project, described below in Chapter 3, indicated that SOLO levels functioned similarly in more advanced mathematics. It is worth noting that, even though the Van Hiele classification would possibly be better adapted to mathematics as done by mathematicians, in the present project the preliminary work made it clear that there was a great need for discrimination at lower levels, and that the highest in both classifications were likely to be irrelevant. An approach similar to that used to define SOLO levels therefore seemed appropriate here, and its working out in the definition of achievement measures for the study is described in Chapter 3.

Connection with attitudes

The research in the second set of studies mentioned above leads one to relate attitudinal and intentional factors to the quality of learning. What the Swedish work had indicated was that intentions were of decisive influence on quality of learning, and that the approach to learning was connected with both motivation and the mechanics of study. The finding of this work were used to illuminate and extend work being done already in Britain and Australia on the connection between achievement, attitudes and study methods among tertiary students.

Early attempts at applying conventional wisdom to the detail of study methods had proved quite unsuccessful, with the predictive power of the use of specific methods being very poor. Brown and Holtzman (1955, 1966) broke away from prescription of detail in designing their Survey of Study Habits and Attitudes. Instead, they used a general description of active and well organised study, and included motivational factors. The work was successful in that it obtained significant correlations between scale scores and achievement, which were predictive beyond the information obtainable from ability tests.

The British work that developed this approach culminated in the large scale and well designed work reported by Entwistle and Wilson (1977), which concentrated on steady, active study methods and intrinsic motivation. What emerged was that it was not so much single factors that predicted achievement, but profiles of attitudes that were best predictive for certain clusters of students, so that multiple paths to success were identified. What is more, the importance of intentions and conceptions of the task were identified, just as in the Swedish work already described.

The resulting next stage of research, which used the Swedish findings to illuminate the idea of the importance of profiles of attitudes, is exemplified in that reported by Ramsden and Entwistle (1981). It produced integrated instruments that identified more global orientations to study, each with its characteristic motivation and strategy, establishing discriminations consistent with dichotomy between deep and surface level processing identified in the Swedish work.

The work of Biggs had also developed through the 1970s, using composite instruments including motivational and personality factors, to examine tertiary students' approaches to study. The work used factor analytic methods in a search for patterns in such approaches, and the results were quite startlingly consonant with the British results. Both sets of studies identified three main approaches among students, called by the British team meaning orientation, reproducing orientation and achieving orientation, which corresponded to Biggs's terminology internalising, reproducing, and organising, respectively.

The most highly valued meaning orientation, often considered as matrix of the highest achievement, is characterised by the intention of deep processing in association with intrinsic motivation. Unfortunately for those wishing to use the idea with mathematics students, both sets of studies operationalise it in terms of wide reading and extension beyond the set task or syllabus, which is more appropriate to study in Arts and Social Science. Work done in Australia by Watkins (1982), using the British instruments, in fact found no evidence of a meaning orientation outside the Arts group, and this is quite possibly due to the instrument being inappropriately formulated for identifying the corresponding approach to study in science and mathematics.

The reproducing orientation is characterised by the intention of surface processing, and motivation to get by and get qualified without taking on more than one has to. Anxiety is a frequent component. Conscripts among mathematics students are very frequent, but surface processing hopelessly ineffective, so such an orientation should be highly predictive of failure.

The achieving orientation is concerned with efficiency, and is characterised by good organisation and a calculating approach rather than intrinsic motivation. It has also been found to predict achievement (for example, by Biggs, 1982). In mathematics, good organisation is generally considered an advantage, but the economy of effort available to the more able student is extreme compared with other subjects, so predictions in this area could be different. What is more, calculating motivation still entails the attempt at understanding, again because mathematics is extremely inaccessible without some degree of understanding, hence requiring some depth of processing.

It is clear that the intention of deep processing is conducive to the integration of knowledge defining higher quality of learning, and that this is essential in solutions to mathematics problems. But it is also clear from the above remarks that the formulation of existing instruments needs modification if the whole approach is to be well adapted to mathematics students, and that it may be hard to discriminate between interest and calculation as motivating factors. It follows that the basic idea of the research, rather than the exact pattern of findings, is what is important to the present study. The later developments intended to produce finer discrimination, such as the work of Entwistle and Waterston (1988). are less relevant. The work of Richardson (1990), which proposes a simple dichotomy between attempts at understanding and a simple intention to get by, appears to be better adapted to the present purpose. It is worth noting also that Meyer and Parsons (1989) present a dichotomous hierarchical solution, obtained by higher order factor analysis.

A guide to the adaptation of instruments for use with mathematics students can be found in the now classic studies of the attitudes and achievement of secondary mathematics students carried out by Fennema and Sherman in the later 1970s and the 1980. From their first large scale studies onward (Fennema & Sherman, 1977; 1978; Sherman & Fennema, 1977), several important points became clear. First, investigation of multiple attitudinal factors was found to be required. Second, attitudes were clearly related to achievement and to the choice of taking more mathematics. Third, it was justified to conclude that the direction of influence was often attitude to achievement rather than the opposite. It is also worth noting that the results obtained by Fennema and Sherman were very clearly supported by the work of Armstrong (1979) using a very large, very carefully selected national sample, as part of the American second National Assessment of Educational Progress of 1978.

The most important factors that were part of students' own approach to studying mathematics were found to be effectance motivation, confidence in learning mathematics, perception of the usefulness of mathematics and attitude to success in mathematics. The latter two seem to be subsumed by choice of degree if one is studying engineering students, but the motivation and confidence factors are highly relevant, and connect with the work on students' learning at tertiary level described above. The effectance motivation studied by Fennema and Sherman consists of a delight in problem solving accompanied by the acceptance of the challenge of mathematics learning. This means that it includes intrinsic motivation and a satisfaction in achievement that could be intrinsic or more success-oriented, which is well suited to handling the difficulty of motivational discrimination pointed out above in connection with tertiary mathematics students. The area of confidence connects with that of anxiety in student learning, and so should be considered, but there are some details that need clarification. First, the target population of engineering students is already upwardly selected for confidence about mathematics learning, so that the factor may have different implications from those it has at secondary level, where variation is very wide. Second, the relationship with anxiety seems more complex than it is at secondary level. Fennema and Sherman (1977) note that their anxiety scale was discarded because of its 90 percent correlation with their confidence scale. Contact with university students, however, and some preliminary piloting for the present study, indicate that there is greater independence of confidence about one's ability to learn mathematics from a factor of anxiety about one's immediate performance in a given subject. This means that upbeat statements of confidence are not necessarily inverse to a sense of anxiety in the obvious way. Given the importance of confidence and anxiety in the work already described, the area should be investigated, but more than one independent aspect may be expected to appear.

One further attitudinal factor that is of special interest in a Keller Plan context, is students' feeling of taking responsibility for their work. In later work by Fennema and other colleagues (Wolleat, Pedro, Becker & Fennema, 1980), students' habitual attributions of achievement to internal, stable factors were found to relate to success in mathematics. The factor of internal versus external locus of control is of great general importance in education, because the Coleman report (Coleman, Campbell, Hobson, McPartland, Mood, Weinfeld & York, 1966) found that internal locus of control was the best predictor of school achievement, rather than background or other attitudinal factors. On the other hand, Watkins (1984) found that Australian students tended to attribute success almost exclusively to hard work, which is an internal but unstable factor. But here, the attribution may be a commonly used culturally acceptable excuse, and so not a good basis for discrimination. What one would like to access, in connection with the Keller Plan, is a factor that resembles internal locus of control in that it represents a commitment to taking action to control one's progress in a subject, but it is not clear that stability is a decisive factor, because it is context that is of interest. The emphasis placed by Bar Tal (1978) on power to influence the causal factor adopted for attribution is perhaps more relevant. As Roueche and Mink (1975, p.4) point out,

> It is important to understand that no-one can try unless he believes he has some chance of success.

It seems more practical not to participate in the general discussion of attributional structures, and instead, merely recognise that the acceptance of responsibility for one's own work is worth investigating in a Keller Plan evaluation.

Separation of students' subject evaluations.

A large number of studies of Keller Plan implementations offer students' enthusiastic subject evaluations as evidence in favour of the plan. It seems here that it is appropriate to separate subject evaluations from the attitudes that form approaches to study. The point is that approaches to study represent a primary field of investigation because of their connection with quality of learning. Subject evaluations, on the other hand, are more like job satisfaction, and form a reserve field of investigation. That is, if some method of instruction is clearly associated with higher quality outcomes or approaches, it is clear that this is more important than students' liking for the method, subject of course to exclusion of unlikely extremes of unhappiness attributable to the method. On the other hand, if there is no clear evidence that methods of instruction differentially foster high quality learning, it is entirely reasonable to discriminate on the grounds of job satisfaction. It is difficult to separate attitudes clearly into one or other category in all circumstances, but extreme differences are obvious, even though factors may in general be part of both areas, as in the case of intrinsic motivation, which is an approach variable, but obviously also an influence on satisfaction.

What follows immediately below is a review of the evidence, in the two areas, separately as far as possible, in studies specifically of Keller Plans in tertiary mathematics.

Studies of attitudes in Keller Plan implementations.

It would be inconsistent to change the inclusion criteria for studies directly concerned with Keller Plans by extending beyond tertiary mathematics. One should, however note that there is a plethora of brief reports on applications of Keller Plans in tertiary science, in which favourable reactions of students are reported, with or without comparisons with conventionally taught classes. Among these, as before, there is the occasional negative result, such as that reported by Roberson and Crowe (1975), and one must also consider the suggestion of novelty effects fading over time, as indicated by Roth (1973) dealing with a variety of methods of individualisation. There are also some overviews at tertiary level, including some accounts of work on mathematics students, that contain results relevant to the present work. In particular, the reviews conducted by Taveggia (1976) and Powell (1980) report evidence that study methods tended to improve in Keller Plan courses, probably because of the clarity of the requirements. Individual relevant studies include that of Flammer (1971), a study of engineering students, which reports an increase in intrinsic motivation in a Keller Plan group, though a formal comparison was lacking. One should also note the evidence offered by Van Damme and Masui (1980), of anxiety reduction in a Keller Plan. In tertiary mathematics there is some material similar to the brief and favourable evaluations, like the conference overview by Sherman, Hassett and Thompson (1978), which reports favourable evaluations without requiring comparisons. In the review that follows, some attempt at evaluating claims, either by comparisons or by links with other factors, is required.

Approach factors.

Results for mathematics implementations are summarised as a guide to discussion, in Table 2.5 Many reports are strongly restricted, which is indicated by 'R' in the table.

Table 2.5. Keller Plans and approach factors.

A. General attitude to mathematics	Report	Control	Comparison results	Changes and correlations
Collard (1989)	R	Achievement	Not significant	Favoured Keller
Edward (1976)		Achievement	Favoured Keller	
Kontogianes (1974)		Not clear	Not significant	Favoured Keller
Locksley (1977)		Not clear	Not significant	
Schielak (1983)		Achievement	Favoured Keller	
Sheehy (1990)		Partial randomisation	Not significant	
Urban (1972)	R	Randomisation	Not significant	

B. Motivation	Report	Control	Comparison results
Pascarella (1977a)		Achievement	Motivation more important in Keller
Sheehy (1990)		Partial randomisation	Keller students reenrolled more
Roberts (1975)	R	Not given	Keller students reenrolled more

C. Study Methods	Report	Control	Comparison results
Stice (1975)	R	Pretest	Favoured Keller

D. Anxiety	Report	Control	Comparison results
Nord (1975)	R	Background	Favoured non Keller
Schielak (1983)	R	Achievement, enrolment	Not significant
Urbatsch (1980)	R	Matching, not clear	Not significant

E. Attribution	Report	Control	Comparison results	Changes and correlations
Roueche, Mink & Abbott (1975)	R	Pretest		Significant links
Urbatsch (1980)	R	Matching, not clear	Interaction: External control worse in Keller	

F. Depth of processing	Report	Control	Comparison results
Rae (1993)		None	Favoured Keller over an unidentified group

What the table indicates is that research is sparse, its quality low, and its reporting often very inadequate. The largest group of studies deals with a global attitude to mathematics, already recognised as unrewarding even at the time of the earlier studies. Hence the generally indeterminate pattern is quite explicable. Overall, there is little that goes against the Keller Plan, but certainly not enough to give adequate backing for a favourable conclusion. In the present context, the area reported on by Rae (1993) is the most relevant, but the work was obviously ancillary to Rae's main project, and the report far from complete enough to allow proper assessment of quality.

Approach factors therefore are not absent from research on Keller Plans, but only just so, and what exists is not very well designed. There is clearly a need for further work.

Subject evaluations

As noted above, subject evaluations are a legitimate area of investigation, and their results may be used as an ancillary decision criterion. That is, extremely unfavourable results could be used in appeal against clear opposing results in primary areas, while any type of definite result could allow a decision if results in primary areas were inconclusive.

Here too, it is a requirement that the studies presented should contain some comparison or evaluation of results, rather than a simple statement that students seemed to like or dislike Keller Plan organisation. Table 2.6 is the summary intended to guide discussion. As before 'R' indicates a restricted report.

Judging by the table, the final appeal to job satisfaction as criterion for a choice between Keller Plans and traditional instruction would reveal fairly clear information. That is, on the whole, Keller Plans receive more favourable ratings, so that if other results gave little reason to choose, one could still discriminate on the basis of subject evaluations. It follows that, although the existing research is fairly consistently favourable, any Keller Plan study would do well to conduct a subject evaluation in the cause of having a reserve criterion available.

Table 2.6. Keller Plans and subject evaluations

Study	Report	Control	Comparison
Brook (1977)	R	Not possible	Anti-handbook ratings favoured Keller
Harris & Liguori (1974)		Achievement	Not significant
Klopfenstein (1977)		Achievement	Not significant
Locksley (1977)	R	Not given	Favoured Keller
Malec (1975)		Not possible	Course rated second in department
Nord (1975)	R	Background	Partly favoured Keller
Pascarella (1977a)		Achievement	Favoured Keller
Thompson (1980)		Randomisation	Against Keller
Urban (1972)	R	Randomisation	Favoured Keller
Van der Klauw & Plomp (1974)		Not clear	Favoured Keller
Watson (1986)		Achievement/ enrolment	Favoured Keller.

Interactions of teaching methods with subgroup membership

In the total group targeted by the present study, there is a natural two-way split based on citizenship and gender. It is claimed here that both the overseas students and the women are specially highly selected groups, so that, although there is a whole body of research on cultural and gender influences on mathematics learning, it is not likely to be a good guide to the formulation of research questions for the study.

Citizenship groups

There is a large minority of overseas students at the University of New South Wales, most of whom are citizens of South East Asian countries. The largest subgroups come from Malaysia and Indonesia, and dominant cultural influences are those of the longer established peoples of the region, together with a strong contribution from the overseas Chinese. There are therefore marked cultural differences between the overseas group and the mass of Australian students, even though, as pointed out already, the division was already blurred because of a long established Australian Chinese community, and is becoming more so because of recent immigration patterns.

But the most important influences on the studies of the overseas group must include their specific position as overseas students, which is likely to represent a selection and stress factor of importance comparable to the cultural differences. First, they are under a very strong obligation to their families, because of the high cost of study in another country, which includes high tuition fees. This combines with the traditionally high value given to diligent study in Asian cultures to make them far more strongly committed than many of the Australian group. Second, they tend to be rather more highly selected for past achievement than the local group, again because of the difficulty of overseas study. Third, a very common path taken by overseas students who want an Australian university place is to get an Australian university entrance qualification after doing a tertiary entrance qualification at home. This means that they are especially well prepared, and older than the Australian group, and so often more sophisticated, both intellectually and in life skills. What is more, the time spent getting an Australian Higher School Certificate has resulted in practice in, and selection by, precisely the life skills needed by students. For such reasons, the group is bound to be highly unusual even within their own culture, making the formation of hypotheses about their reactions to different teaching methods difficult. The possible interaction with language skills is a further complicating factor.

It seems plausible, therefore, that the overseas students' position as overseas students is a factor at least as important as culture. In any case, an assumption of homogeneity of culture in the region is not necessarily justified. Apart from background knowledge about more subtle distinctions, there is an obvious split between the indigenous communities of South East Asia and the overseas Chinese, and there is research evidence of attitudinal differences between national communities For example, Swetz, Langguling and Johar (1983), found that Malaysian students' attitudes to mathematics were dissimilar to those of Indonesian students.

It seems to follow that it is highly necessary to retain the separation of the overseas students in analyses, but that it is not useful to try to predict from the cultural differences alone, or even from these differences in conjunction with situational factors. For example, one might certainly hypothesise that culture and situation would imply more diligent study methods among overseas students, but there is no reason to predict greater advantage for Keller Plan teaching. What is more, it is hard to say how the language factor would operate, without a special investigation of questions like the relative advantages to such a group of overseas students that might result from written as opposed to spoken presentation of material, and whether this makes much difference in mathematics, given the common notation and the tradition of teachers writing full notes on board or projector in lectures. For present purposes, what this means is that it is not possible to take on investigations of the scope and detail needed for prediction. The overseas group will therefore be compared with the Australian group, but the intention is to be exploratory.

Gender

The target group contains a subgroup of women of non-negligible size, at least among the chemical engineering students (41% of the achievement sample), because for some time chemical engineering has been the engineering degree most frequently chosen by women students at the University of New South Wales, even though there are still considerably more men in the course. There is also a large body of research on gender and mathematics learning that has established the existence of gender related differences in approaches to study. In particular, the studies by Fennema and Sherman already referred to (for example, Fennema & Sherman, 1977, 1978; Sherman & Fennema, 1977) indicated that males tended to be more confident about learning mathematics and more aware of its vocational importance. There is also related work on attributional styles (Wolleat et al., 1980) that found that males tended to have more optimistic attributional styles about learning mathematics than females, which are likely to make the males more active and persistent in their mathematics work. In research on approaches to study at tertiary level, there is also sometimes evidence for gender-related differences, such as those found by Meyer, Dunne and Richardson (1994).

The problem with studies of gender and mathematics is that they are most often concerned with whole populations, and the problem with studies of approaches is, as argued above, that they are often not well adapted to mathematics learning. That is, one may use findings on mathematics learning in general as a guide to reformulating approach factors in a way more suited to mathematics, but the male and female groups targetted in the present work are unusual, and probably not even selected equivalently themselves, because of the different levels of commitment required of males and females in choosing to study engineering. That is, both male and female groups consist of upwardly selected survivors in mathematics, but such selection is likely to be more stringent in the case of the women. It follows that it is not justified to expect gender differences similar to those found in either set of samples, that is, those that are inclusive either by target population or by university subject. One might just as reasonably predict gender related differences opposite to those found, or possibly no differences. It seems wiser therefore explicitly to adopt agnosticism about predicted outcomes, but to perform appropriate gender comparisons in the areas of achievement and attitudes.

Implications

The research review indicates that it is possible to define suitable achievement scores for comparisons between groups being taught in traditional and Keller Plan organisation. The detail of score construction is in Chapter 3.

The potential importance of attitudinal factors has also been made clear. The use of questionnaires as assessment instruments is nothing new, but the adaptation of ideas and insights from existing work to the specific context being studied is still required. Details of instruments for assessing attitudes are also presented in Chapter 3.

Chapter 3. Method

Introduction

The chapter covers samples and their subgroups, the teaching methods used with different groups, procedures and instruments, and the method used to define achievement scores. Most of the work was planned for quantitative analyses, but it is supplemented by a small amount of interview material suitable mainly for qualitative analysis.

Samples

The samples were drawn from two sets of engineering students studying mathematics subjects designed for engineers, at the University of New South Wales. The three groups involved had experienced either Keller Plan or traditional organisation of teaching in their prescribed Year 2 mathematics subject. Details of the teaching methods are given below.

The three groups from which samples were drawn were as follows.

Group 1: Chem93. Chemical engineering students, doing second year mathematics in 1993, taught in a newly instituted Keller Plan implementation. At third year level, they did a further mathematics subject, taught in a long established Keller Plan implementation, that described by Barrett and Prokhovnik (1980).

Group 2: Civ93. Engineering students from other branches of engineering, mainly civil, mechanical, mining or surveying, doing a common subject taught by traditional methods.

Group 3: Chem92. The chemical engineering students who did Year 2 mathematics in 1992, using the same syllabus taught to Chem93, but taught by traditional methods. They too continued into the third year Keller Plan mathematics subject mentioned above.

For each group, attitudinal data were collected in class, whereas achievement material was available for whole enrolment groups. Maximal available samples were used in each analysis, so that sample numbers for achievement results are larger than those for attitudinal results, and numbers fluctuate slightly for larger than those for attitudinal results, and numbers fluctuate slightly for variables of the same basic type. The interview sample was very small, consisting of nine Keller Plan students, and seven from the traditionally taught group.

It was noted in Chapter 2 that cultural differences might appear, and that the large minority of overseas students at the University of New South Wales not only come from a different culture, but, as full fee paying students living away from home, are under very strong pressure to succeed. The sample is therefore usually broken down by citizenship, into subgroups labelled Australian and overseas, where the overseas group consists overwhelmingly of citizens of countries in South East Asia. One should note, however, that actual cultural differences have been somewhat blurred by immigration, and the Australian group contained Asian immigrants, some of quite recent arrival.

Samples and analyses are also usually broken down by gender.

Overall numbers are in Tables 3.1 and 3.2.

Table 3.1 Achievement sample

		Australians		Overseas	
Year 2	Group	Women	Men	Women	Men
	Chem93	42	97	17	21
	Civ93	44	186	5	35
	Chem92	35	104	11	17

		Australians		Overseas	
Year 3	Group	Women	Men	Women	Men
year 3	Chem93	13	36	10	13
	Chem92	15	40	4	10

Table 3.2 Attitudinal samples

First data collection

		Australians		Overseas	
First	Group	Women	Men	Women	Men
Start Year 2	Chem93	43	86	15	17
Start Year 2	Civ93	41	172	5	35
End Year 2	Chem92	19	51	4	10

Second data collection

		Australians		Overseas	
Second	Group	Women	Men	Women	Men
End Year 2	Chem93	39	83	14	19
End Year 2	Civ93	30	121	3	26
End Year 3	Chem92	14	30	3	26

Longitudinal data available

		Australians		Overseas	
Longitudinal	Group	Women	Men	Women	Men
	Chem93	34	63	14	17
	Civ93	26	94	2	25
	Chem92	13	28	3	5

The teaching methods

The groups Civ93 and Chem92 were taught in the way traditional in Australian universities. That is, they received lectures in large groups, supplemented by tutorials in small groups and, in the case of Civ93, also supplemented by large group problem solving demonstration classes.

The group Chem93 was taught in a new implementation of the Keller Plan, which was, however, modelled on the very long running Keller Plan implementation taken by the third year chemical engineering students (see Barrett & Prokhovnik, 1980), and was taught by many of the people who had experience of the third year subject. This means that smooth running of the new implementation was much facilitated. It is possible that novelty effects could have been produced, but it is hoped that the collection of data from all groups indicated interest in them, which affords some control for Hawthorne effects.

The mathematics syllabus for the two chemical engineering groups was the same. The syllabus studied by the other engineering students was very similar, but somewhat longer. The Keller group worked individually from printed notes and sets of problems. They used class time in even numbered weeks for consultation with teaching staff, if they wanted it. In odd numbered weeks, they were given the opportunity to take a test, which was marked immediately by the staff who would otherwise have taught the tutorials in the subject. Discussion and explanation in conjunction with marking were possible, and students had the chance to defend and improve their answers. There was no final examination. A pass in the subject was given for demonstrated mastery of a minimal core of topics, and higher grades could be earned by completing additional topics.

The material was not at a very high level of mathematical sophistication, but was much more advanced than that of many American implementations. The level of material and the context implied the necessity of adapting certain features of the Keller Plan away from the detail of the classic form described by Keller (1968).

Length of units has been mentioned above. It is recalled that the minimal comprehensible unit in mathematics is longer and more complex than the programmed learning ancestry of the plan might seem to require, but that adjustment to fit a particular type of material was recognised in later formulations by Keller himself (Keller & Sherman, 1974). A detailed discussion of unit length is to be found in the work of Barrett and Prokhovnik (1980) in connection with the established third year subject on which the present implementation was modelled.

The assessors of test work were staff, rather than senior students, because of the relatively high level of the material, and also because the constraints of class time required experienced markers who could do the work fast and consistently. Since staff-student relations in the subject were informal and friendly, the greater distance between assessor and assessed is believed not to have caused difficulties.

There was also a certain amount of forcing of the pace, to fit the subject inside the traditional academic year. This sort of provision, however, has been so common even from early days as to be regarded as normal (see the reviews by Boud et al., 1975, and Imrie et al., 1980).

The essence of the plan that remains consists of self-organisation, multiple test opportunities, and the mastery requirement. The meaning of the term "mastery" is a perpetual discussion point in reviews of studies and practice (see, for example, Imrie et al., 1980, p. 107), but in the present case is defined by practice, justified by the ease of detecting contradictions in mathematical arguments. Students demonstrated mastery of a topic by solving a set of problems on it, correctly except for minor errors not involving inconsistencies in the main argument. Both informal checking among markers, and the re-examination of marked scripts involved in the construction (described later in this chapter) of achievement scores for the present project, indicated that common understanding and application of the criterion by markers were very consistently demonstrated.

Data collected and procedures used

Three sets of data were collected for the main part of the study, and the information was supplemented by a small number of interviews.

The main study collected background information, achievement scores and attitude scores. Background information consisted of name, age, gender and citizenship.

Achievement material came from university records and the test and examination scripts produced by the sample for assessment in mathematics in Years 2 and 3. The examination scores from the common Year 1 mathematics subject were used as a covariate, for statistical control of initial differences. Achievement scores for subsequent mathematics subjects were constructed using direct re-examination of the students' work done for assessment. Justifications and details of score construction methods are to be found later in the present chapter.

Attitude scores were obtained from questionnaires about approaches to study, and also about students' perceptions of the main mathematics subjects in which they were involved. The attitude questionnaire consisted of 32 items, each scored on a five-point scale. The questions are reproduced in Appendix 2, and the formation of attitude scales is described later in this chapter. Attitude questionnaires were collected in class from Chem93 and Civ93 at the beginning of Year 2, and from all three groups at the end of Year 2. Students were very cooperative.

Interviews were semi-structured, using a fixed set of questions, but allowing open ended answers and requests for clarification and detail. The interview schedule is to be found in Appendix 3. Students' responses were tape recorded, with their consent.

Achievement scores

Examination marks as covariate

Since it was intended to evaluate effects of Keller Plan teaching, it was obviously desirable to control for initial differences of mathematics knowledge and previous performance in mathematics.

All engineering degrees at the University of New South Wales have a common Mathematics 1 requirement, so that all the groups entering Year 2 tend to have at least this subject or an equivalent as common background in mathematics. The first year subject is taught at two levels, but the syllabuses are not very different, there are common questions in examinations, and marks for both levels are considered together as a single distribution. The higher level is taught at rather greater depth, for the benefit of more able and more interested students.

It was therefore considered justifiable to use the examination marks from Mathematics 1 as a control variable. The remarks above indicate that the existence of two levels need not present a problem. It was also considered reasonable to treat marks from adjacent years as roughly equivalent in meaning. Distributions of marks are very stable from year to year, and there are strong reasons for regarding the mark levels as reasonably stable in meaning. That is, the justification here is that the marks come from very large groups working on very stable syllabuses, and whose general level of selection remains similar from year to year. Thus Year 1 examination marks seemed to be a very good control variable.

An additional controlling check was made by comparing the first year examination marks of the two groups entering Year 2, that is, Chem93 and Civ93. A factorial analysis of variance by group, gender and citizenship, reported in detail in Appendix 1, indicated that there were no significant differences between the teaching groups involved in the present project. There were also no differences between citizenship groups, but there were gender differences, because the women in the sample had significantly higher scores than the men. No interactions were significant.

Year 2 and Year 3 achievement scores

The achievement scores used in comparisons were planned to represent quality of learning, using a division of the written answers to assessment tasks into levels similar to those defined by the SOLO (Structure of Observed Learning Outcome) taxonomy (Biggs & Collis, 1982), in which it is the degree of integration of relevant knowledge that counts. The origins of the taxonomy and details of its definition are described above in Chapter 2. What is presented here is a brief outline of arguments intended to justify using SOLO levels as a guide to classifying solutions to mathematical questions, and a description of the application of the classification to the results for some typical tasks. Further documentation and examples are to be found in Appendix 1.

As described above, the SOLO taxonomy was first presented by Biggs and Collis (1982) in an application to single tasks at upper primary to early secondary levels. The work of Chick et al. (1988) found it to be an adequate guide to the classification of individual mathematics errors in the work of first year trainee teachers, and the report of the work gives good documentation to support the argument. There is therefore some justification for applying it to short mathematical tasks at a somewhat higher level than that originally presented. It is argued here that there are features of the present project that make the classification task simpler, and thus more tractable, than it would be in a more general context.

To begin with, the important relevant issue raised in the literature is precisely that of the generality of application of classifications such as the SOLO taxonomy. In the present project, it is essential that classification should be validated across different tasks, but it is also clear that the differences involved are relatively small, and do not involve anything like the difficulty of the question of widely varying tasks in fuller generality. That is, the justification of the classification is considerably simplified, because the tasks involved are mathematical, and hence constrained by the convergence of solutions, and made highly comparable by their reference to very similar syllabuses, examined at a similar level of difficulty. It is claimed that one can support the claim that levels of difficulty are similar, by appealing to professional consensus among teachers, whose judgment is continually being refined by evidence from students' actual handling of tasks. One can certainly check directly that there is consensus, and this was checked for the present project by an informal survey of the staff who taught the groups in the sample. The validity of the consensus is based on the fact that the groups of students being taught do not vary markedly from year to year, so that professional judgment of difficulty in terms of the task as criterion is perpetually being checked against performance norms in similar groups. In the mathematics subjects done by the present sample, stability of syllabuses and examining styles is clear, and the groups of students are quite large, and selected by criteria that do not vary sharply from year to year. That is, institutional and selection inertia give a stable validation group for judgments about the difficulty of tasks.

The second issue that must be dealt with is the validity of classification of students' response to assessment demands into categories that may be regarded as ranked with respect to quality. When there is a single mathematical task, the ranking of outcomes is made by traditional marking of answers. It is quite easy to support a claim that such methods can be highly reliable: that is, that the notorious unreliability of examinations can be overcome when it is mathematics that is being examined. First, errors in mathematical reasoning can be demonstrated, quite easily when the argument is neither long nor complex. Second, one can appeal to experience in marking large groups such as university first year classes, or candidates in public examinations. The first thing to note is that students' errors tend not to be very original. That is, when a solution contains errors, it tends to fall into one of a small number of alternative patterns. After the first fifty scripts, it is almost unheard of to encounter anything really new. This means that the universe of outcomes presented by the students is quite small, so that each pattern can be discussed to produce a set of marking decisions that can be consistently implemented. In the marking of very large groups done at the University of New South Wales, checking procedures have been known to find, after the initial discussion and comparison stage, that it is extremely rare to find more than about one percent disagreement, and almost unknown for assessments to differ by more than one mark in ten. Reliability can therefore be achieved, with some effort.

It remains to validate marking procedures. The linearity and convergence of mathematical arguments help here. That is, on a short mathematical task, it is not difficult to identify stages in alternative arguments, and to rank these stages by how much of a full set of required steps they achieve correctly. This is helped by the tendency noted above for solutions to cluster into a small number of patterns. Validation of the marking of a short mathematical task in terms of ranking by completeness of argument is therefore quite easy to defend. The tasks relevant to the present project are quite short, and so their marking is accessible to such validation.

Requirements for the project are, however, more extensive, because one needs a classification of outcomes that can extend across different tasks. It is claimed that one can use the SOLO taxonomy as a guide to the creation of a transferable procedure that is similar to normal examination marking, given that requirements are simplified in the ways described above. One should point out further that, within the project, the work produced by students is a finite universe, in which ranking can be done without one having to claim open ended generality. The advantage of the SOLO taxonomy is that it proposes a fairly coarse ranking, and gives easily transferable criteria for the assignment of ranks. The basic divisions of the taxonomy, in contexts like the present, result in practice in a three level assignment task. This is because the highest level, extended abstract, is well beyond what could reasonably be found in not very advanced engineering mathematics subjects, while the lowest category, prestructural, equivalent to producing nothing that is right, is so obvious as to present no classification problem. The next lowest, unistructural, in which only one idea is present, is also easily operationalised, and familiar to the experienced marker. This leaves one discrimination, and the SOLO criterion for making it is clear. That is, relational arguments do not contain contradictions, whereas multistructural ones do (Biggs & Collis, 1982, pp. 24-25). This is particularly convenient because of its consonance with decisions in enforcing the Keller Plan mastery requirement: to show mastery of a topic, the student's work should be free of contradictions. That is, in the implementations of the plan under consideration, a student who makes an error showing lack of understanding does not pass the test. Lack of understanding here means the production of false conclusions that go beyond small slips, in arithmetic or non-essential detail. That is, fatal errors are those that alter the problem significantly, use underlying theory invalidly, or regress to a simpler situation when more developed methods are required. The classification includes leaving out essential portions of argument, and, a fortiori, question begging and mistaken directions of implication. If a small error distorts the problem to an extent that makes a coherent solution impossible, this is also considered fatal.

In the course of reviewing test and examination scripts for the present project, it was quite easy to apply the distinction separating relational from multistructural, but a distinction within the relational category, between flawed and perfect solutions, seemed appropriate. The implied problem remaining is that the definition of the multistructural category is open to the criticism that it operates purely by complementation. That is, one must be able to establish that it is not just a ragbag containing everything that is not obvious. But the prescriptive definition of the category, as containing those solutions that contain more than one relevant idea, but are invalidated by contradiction, is actually clear enough for definition to operate without pure resort to complementation, so that the total classification is intrinsically defined. But even so, the range of cases it permits is quite large. Here one may usefully appeal back to normal marking methods, which discriminate on the basis of ranking by completeness of argument. The fine discriminations that can be made within one task may not be available, but a rough split into more and less complete arguments is quite easy to apply consistently. It is therefore claimed that a rough but reasonably valid classification of solutions to short mathematical problems at a similar level can be obtained. The final justification, however, requires some direct supporting evidence.

First, it may be noted that pilot work for the present project afforded some evidence that a classification of the type required could be made reliably. A set of several hundred answers to linear algebra questions was classified into six levels, according to the criteria described immediately above. After an interval of two months, classification of the set was done again, and the rate of consistency was found to be about 98%.

The validity requirement could be supported by correlations of rankings with examination scores obtained by normal marking of the same material, and these were in fact high. But the classification was intended to extend ranking beyond normal marking, so it seems appropriate instead to supply some evidence of how the classifications were actually made, using material from the project. What follows here is an account of the process, as applied to answers to two second year examination questions, found in a random sample of fifty scripts from Year 2 Chemical Engineering students. Full sets of examples of student's work at each level are in Appendix 1.

Question1

Find the general solution to the following differential equation:

$$y' - \frac{x^2}{x^3+1}y^{37} = 0$$

Solution

Separate the variables to get

$$y^{-37}dy = \frac{x^2}{x^3+1}dx \qquad (1)$$

$$\int y^{-37}dy = \int \frac{x^2}{x^3+1}dx \qquad (2)$$

$$-\frac{1}{36}y^{-36} = \frac{1}{3}\ln|x^3+1| + c$$

(3) (4)

Classification of answers

Prestructural		Nothing right
Unistructural		Only as far as (1)
Multistructural	Lower	Correct to (2) but no correct integration
	Upper	Correct to (2), plus either (3) or a near correct (4)
Relational	Lower	Correct up to (4) except for slips in arithmetic or no absolute value signs
	Upper	All correct

Patterns obtained on fifty scripts

Prestructural		1:	No separation of variables, no other (wrong) classification of the differential equation
Unistructural		2	Classifies the equation as linear, and persists
		1:	States separable but no separation
		2:	Separates but does not integrate
Multistructural	Lower	2:	Separates, tries integration with no success
	Upper	6:	Separates, integrates to get right y-power only
		6	Separates, approximately right logarithm, and modulus signs or constant missing
Relational	Lower	15	Miss modulus signs
		2	Method valid but wrong coefficient for logarithm
	Upper	13	Complete correct solution

Question 2

Let y(t) be the solution of the initial-value problem

$$y'' - 3y' + 2y = r(t), \qquad y(0) = 0 \quad y'(0) = 1$$

where r(t) = u(t) u(1-t) and u(t) is the Heaviside step function.

(a) Sketch the graph of r(t).

(b) Show that the Laplace transform of y(t) is

$$Y(s) = \frac{1 - e^{-s}}{s(s-1)(s-2)} + \frac{1}{(s-1)(s-2)}.$$

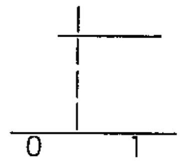
Indicate the term corresponding to the initial condition.

(c) Find the inverse Laplace transform of just the first term in $Y(s)$, that is, find

$$L^{-1}\left\{\frac{-e^{-s}}{s(s-1)(s-2)}\right\}$$

Solution

(a) Graph of $r(t) = \begin{cases} 0 & t<0 \\ 1 & 0<t<1 \\ 0 & t>1 \end{cases}$



(b) Taking Laplace transforms

$$\begin{aligned} s^2Y - 0\cdot s - 1 - 3(sY-0) + 2Y &= \int_0^\infty e^{-st} r(t)dt \\ &= \int_0^1 e^{-st}dt = \left[\frac{e^{-st}}{-s}\right]_0^1 = \frac{1-e^{-s}}{s} \\ (s^2-3s+2)Y &= \frac{1-e^{-s}}{s} + 1 \\ Y &= \frac{1-e^{-s}}{s(s-1)(s-2)} + \frac{1}{(s-1)(s-2)} \end{aligned}$$

in which the second term on the right corresponds to the initial condition.

(c) $\dfrac{1}{s(s-1)(s-2)} = \dfrac{\frac{1}{2}}{s} - \dfrac{1}{s-1} + \dfrac{\frac{1}{2}}{s-2}$

So the required inverse Laplace transform is

$$\frac{1}{2} - e^t + \frac{1}{2}e^{2t} - u(t-1)\left\{\frac{1}{2} - e^{t-1} + \frac{1}{2}e^{t-2}\right\}$$

Classification of answers

Prestructural		Nothing right
Unistructural		One correct fact, such as knowing the definition of the Heaviside function, and nothing else
Multistructural		Any answer that does not get r(t) right, or does not provide any argument for getting its Laplace transform, shows clear lack of understanding, and so falls into the multistructural class or lower. So does any answer failing to find the inverse transform in part (c), unless errors are only minor slips. A minimum requirement is to get the Laplace transform of the left hand side of the equation approximately right.
	Lower	Something wrong with r(t) or its transform, and the inverse calculation wrongly formulated
	Upper	Require one of the above right
Relational	Lower	Correct with slips. Accept also those who apparently overlook identification of the term corresponding to the initial condition.
	Upper	All correct

Patterns obtained on fifty scripts

Prestructural		8	Nothing correct
Unistructural		2	Heaviside only
Multistructural	lower	16	Left hand side right, r(t) or its transform wrong, wrong method for inverse transform
	upper	9	r(t) and its transform wrong, rest right
		4	Transform of r(t) unjustified, no identification, inverse slightly wrong
		5	r(t) and its transform right but inverse missing or done by wrong method
Relational	lower	2	Slip in arithmetic
		5	No identification of term associated with initial condition
	upper	1	Correct

Conclusion

The above gives an illustration of the splitting of responses into broad levels, performed on two tasks of different length and complexity, but using similar rules, based on the completeness and consistency of the written work. It is claimed that such rules afford a ranking of outcomes that is similar to that obtained in normal examination marking, and uses the same detailed work, thus having access to more information than would be obtainable from other types of test, notably those using multiple choice questions.

Completeness of material

The achievement scores were calculated from the scripts produced by students in examinations and tests. The work considered was the best the student could produce under test conditions, on the whole syllabus in their Year 2 mathematics subject. That is, for traditionally taught students, all the examination scripts for the year were used. For Keller Plan students, the best test on each topic was scored, and if no test was taken, the score given for that topic was zero. This means that students not completing the Keller Plan subject were included in the sample, and given zero for topics not assessed. Such a procedure serves two purposes. First, it means that both groups were assessed on whole syllabuses, so that their learning over the whole year in a complete set of topics is being compared, which is an obvious and essential point. Second, it works also as a precaution against the danger of assessing the Keller Plan on a sample biased by the dropping out of students who are not succeeding. If anything, it biases in favour of the traditionally taught, because those dropping their mathematics subject before the half yearly examination would not appear in the achievement sample, while Keller dropouts from the first half of the year would appear, and would get low overall achievement scores.

Attitude scores

Information about students' attitudes to their work, and their perceptions of their mathematics subject, was collected by questionnaire. The instrument consisted of a thirty-two item questionnaire, for whose items agreement or disagreement was indicated on a five point scale. Twenty-four items were about approaches to study, and eight about perceptions of the subject. The questionnaire is reproduced in Appendix 2.

The items were selected in the light of the research on attitudes and approaches to study described in the literature survey. Items were therefore initially interpretable in accordance with overt content and with results from other studies that used similar items.

Attitude questionnaires were collected in class from all groups, at the first class of the year. These gave attitudes before second year experience for the groups

Chem93 and Civ93, and after such experience for Chem92. Later in the year, answers to the same questions were collected again, in lecture classes for the traditionally taught, and, for Keller Plan students, either after the last compulsory test, or, in the case of those who intended to drop the subject without failure, after the last test they took.

Students were cooperative, and seemed quite happy to identify themselves on answer sheets, although comments on the subject were rather more often omitted, indicating some greater unwillingness to expose themselves in this area.

Initial analyses were done using data from the first set of questionnaires, on a total sample of 488 students. Factor analyses were performed on two separate sets of items, those dealing with approaches to study as one block, and those dealing with subject perceptions as the other. The factor analyses were done using the SPSS programme FACTOR (Norusis, 1990), with factors extracted by principal components analysis, and rotation of axes done by the Varimax method. Because the number in the sample was large, quite small loadings were significant. For this reason, the conventional level of absolute value at least .3, rather than statistical significance (Gorsuch, 1983, p. 210), was adopted as the criterion for a loading to be considered salient, and only such loadings are recorded in the results presented here.

Students' approaches to study

Five approach scales were originally planned, dealing with intrinsic motivation, active and diligent study methods, confidence, anxiety, and control over one's own studies. Items were scored on a five-point scale, with mixed positive and negative wording, and conversion of scores to give high scores to attitudes assumed to be favourable. Pilot work, however, indicated that some control items needed reassignment, and it was decided to discard some and absorb others into different areas, leaving just one direct question about feeling in control of one's own work. The rest of the items were factor analysed, and results were assessed using the criterion that a factor would be used if it was associated with an eigenvalue greater than 1. Four factors were expected, because there were four areas of emphasis in the content of the surviving items. Pilot work had indicated that the upward and downward aspects of confidence and anxiety might usefully be kept separate. Four factors did emerge. Eigenvalues are reported in Table 3.3, together with the percentage of the variance the corresponding eigenvector accounts for, which is the same before and after rotation when the present method is used (Norusis, 1990, p. 324). The rotated factor matrix is in Table 3.4.

Table 3.3. Factors and eigenvalues

Factor	Eigenvalue	Percent of variance
1	5.64	25.6
2	3.15	14.3
3	1.42	6.4
4	1.30	5.9

The rotated factors are mostly quite easy to interpret. Factor 1 is almost entirely made up of items dealing with confidence, and is of dominant importance in the amount of variance it accounts for (26%, nearly twice that of the next factor). Such a pattern is compatible with the work of Fennema (1985) in which the main argument is that confidence is of crucial importance in the development of autonomy in mathematics learning, and autonomy is a key factor in achievement. The present sample, being already selected for success, is likely to have high confidence about mathematics, compared with the general population.

Factors 2 and 3 relate to items originally intended to measure intrinsic motivation and diligent study methods. Two items originally intended to reflect motivation (Items 18, 23) clearly associate more strongly with the study methods group. Because these deal, respectively, with the importance of doing well and the intention of trying hard to understand the work, there is no problem about reassigning these to a scale still called Study methods. There is some overlap between the two factors, given that two items load similarly on the two factors (Items 2, 21). But there appears to be no problem here either, for two reasons. First, interpretability rather than total separation of factors was being sought. Second, both items loaded more heavily still on Factor 4, which was dominated by items whose positive direction indicated low anxiety. The two items under consideration reflect ease in getting organised and settling down to work. Their content thus involves the absence of unreasonable problems about the mechanics of study, so that they can be grouped with items about anxiety on the grounds that they reflect reduced worry about doing required mathematics because of a systematic approach that accompanies the absence of irrational fear.

There is also overlap between Factors 1 and 4, involving some confidence items, but the difference between the factors is clearly established, and, in all cases but one, items can be allocated to one or the other on the basis of relative loading sizes.

Table 3.4 Factor loadings

Area	Item	Factor 1	Factor 2	Factor 3	Factor 4
Confidence	3	.80			
	9	.46			.38
	11	.76			
	15	.57			.41
	17	.75			
	19	.45			.44
Study methods	2		.33	.34	.49
	7		.52	.31	
	12		.31		
	16		.68		
	21		.39	.37	.49
	24		.80		
Motivation	1			.70	
	6			.75	
	10			.76	
	14			.66	
	18		.56		
	23		.67		
Anxiety	5	.46			.63
	8				.62
	20				.60
	22				.54

It seems reasonable, therefore, to use four scales, and the separate item (number 13) about control of one's own work, with items distributed as follows.

Table 3.5 Scales

Name	Items
Confidence	3, 9, 11, 15, 17, 19
Study methods	7, 12, 16, 18, 23, 24
Motivation	1, 6, 10, 14
Anxiety	2, 5, 8, 20, 21, 22
Control	13

The reliabilities for the scales, using Cronbach's alpha, are as follows.

Table 3.6 Reliabilities

Scale	1	2	3	4	Total
Alpha	.82	.71	.78	.68	.84

The coefficients are well above the limits considered satisfactory for research (.5 to .6, see Nunnally, 1967, p. 226).

Subject evaluation

The eight items (25-32) involving students' perceptions of their mathematics subject afforded just one factor. The scale consisting of the items had a reliability coefficient of .82 (using Cronbach's alpha again).

Analyses of scores

Attitude scores were correlated with Year 1 examination marks, as a check on the justifiability of considering high scores to be associated with good academic performance as well as aspects of job satisfaction. Correlational links with the achievement scores calculated for later years were also examined, for similar reasons, as well as for comparison with the first year links.

The attitude and achievement scores of groups taught in Year 2 by different methods were compared using factorial analysis of covariance controlled for Year 1 examination marks. The natural factoring by teaching group, gender and citizenship was made difficult because some cells were very small. The problem was resolved in context, and solutions are reported in Chapters 4 and 5.

The Year 3 achievement of the groups Chem92 and Chem93 was followed up as an extension of the main design. Permissible conclusions from the Year 3 results are necessarily restricted. This issue is discussed in context.

The interview substudy

A very small sample of students from Keller Plan and traditionally taught groups also took part in interviews about their experience in Year 2 mathematics subjects They were invited individually. The questions dealt with the same areas as the questionnaires, that is, approaches to study and subject perceptions, and were intended to elicit points difficult to access in closed form, or salient to students but not asked about elsewhere. The material was used to illuminate the quantitative results.

Ordering of results.

Results are presented in Chapters 4-6. Attitude questionnaires are described first, in Chapter 4, because some analyses of the scales are required before one can assess relationships between attitudes and achievement results. The achievement results are in Chapter 5, and the interview results are given last, in Chapter 6, because they are more illuminating when one has the other results available.

Chapter 4. Attitudes: results

Introduction

The chapter opens with correlations between attitude scales and previous examination marks, calculated to confirm the implicitly postulated favourability of the attitudes being assessed. Comparisons between teaching groups, factored by citizenship, are then examined. Because preliminary analyses indicated no gender related differences, and using gender as a third factor tended to make some sample divisions quite small, comparisons involving gender are reported only in Appendix 2. Changes in attitude scores over Year 2 are then examined for significant differences from zero and for group differences.

Correlations with achievement

Approaches to study

The overt content of the items in the attitude scales was chosen, in the light of previous research, to address questions where implications for job satisfaction were obvious, but which, it was inferred from previous work, were also important to the quality of learning outcomes. The inference remains to be tested directly. To do this, the first set of attitude scores, which referred to experience in Year 1 mathematics, were correlated with Year 1 examination marks for the whole sample. Correlation coefficients are in Table 4.1.

Table 4.1 Total sample: correlations among attitude and Year 1 examination scores

	Study Methods	Motivation	Anxiety	Year 1 exam
Confidence	21**	40**	.58**	.29**
Study methods		.45**	.09	.13**
Motivation			.29**	.14**
Anxiety				.24**
Max N = 488	** p < .01			

The control item

The item stating that students felt they were in control over their mathematics studies was also correlated with the scales and the achievement scores. Results are in Table 4.2

Table 4.2. Correlations with control

Confidence	Study Methods	Motivation	Anxiety	Year 1 exam
.20**	.36**	.29**	.22**	.11*
Max N = 488	* p < .05	** p < .01		

Perceptions of Year 1 mathematics

The eight items (numbers 25-32) involving students' perceptions of the first year mathematics subject afforded just one factor, and the resulting scale had a reliability coefficient of .82 (Cronbach's alpha), and correlated .20 ($p < .01$) with the Year 1 examination mark.

Conclusion

The evidence cited immediately above is compatible with the argument that the attitudinal scores represent factors positively associated with both job satisfaction and achievement. One may therefore assume that, in comparisons, higher scores represent some advantage.

Group comparisons at the end of Year 2

Separate analyses were conducted for each attitude area, because it was desired to investigate the patterns emerging in each separate area as well as the general question of group differences.

The natural analyses that suggested themselves were factorial analyses of variance by teaching group, gender and citizenship. Unfortunately, some of the cell sizes produced by such analyses were very small. The compromise adopted was as follows. Three-factor comparisons were performed with the non Keller Plan groups combined. Details are in Appendix 2. These produced only one significant result involving gender; a significant interaction between teaching group and study methods, indicating that the Keller Plan was more favourable to women's study methods than to men's The main comparisons were therefore made by two way factorial analyses of covariance by teaching group and citizenship group, with control for the Year 1 examination mark. The control for Year 1 achievement serves the purpose of making the attitudinal comparisons somewhat independent of initial achievement differences.

A full set of analyses is presented in Appendix 2. Results are summarised below, in Tables 4.3 and 4.4.

Table 4.3. Teaching group comparisons

	Means			Significance of F ratio
	Keller	Non-Keller		
Variable	Chem93	Civ93	Chem92	
Confidence	3.34	3.37	3.52	p < .05
Study methods	3.54	3.32	3.08	p < .001
Motivation	3.28	2. 96	2.96	p < .001
Anxiety	2.92	3.16	3.37	p < .05
Control	3.96	3.26	3.15	p < .001
Subject	3.25	3.26	3.15	ns
N = 369				

The Keller group's study methods, motivation and control scores are thus significantly higher than those of the two non-Keller groups, while the reverse is true of confidence and anxiety scores. No differences in subject perceptions were found. It follows that there is a trade off between favourable attitudinal factors and teaching method, and the uncertainty is not reduced by the subject perception scale results. Conclusions can be reached, therefore, only when the full set of results has been presented.

Table 4.4. Comparisons between citizenship groups

	Means		Significance of F ratio
	Australians	Overseas	
Variable			
Confidence	3.38	3.41	ns
Study methods	3.31	3.52	p < .01
Motivation	3.01	3.34	p < .001
Anxiety	3.16	2.97	p < .05
Control	3.76	3.89	ns
Subject	3.25	3.15	ns
N . 369			

No interactions were significant, but the subject evaluation variable showed a near-significant (p = .07) interaction between teaching group and citizenship group. This reflected a pattern of means indicating that the overseas group tended to express greater preference for the Keller Plan than the Australian group.

As one might expect from what was outlined above about cultural differences and the pressure on overseas students to succeed, the overseas group had higher motivation and study methods scores. Their lower anxiety scores (indicating greater anxiety) are perhaps puzzling, particularly in view of the lack of difference in the more upbeat expression of confidence. The explanation that suggests itself is that they are indeed confident about their competence, but regard performance as the result of taking thought and taking pains. More careful thought about what was required of them would therefore explain their higher anxiety scores. As before, further discussion in the light of the full set of results is required.

The near significant interaction between citizenship and teaching group admits several explanations, of which the most plausible is the elimination of the particular difficulties with spoken language that usually arise in large group lectures, which are not part of Keller Plan teaching.

Change scores

Attitude scores collected at the beginning, and towards the end, of second year mathematics studies were available for the two groups Chem93 and Civ93. Changes in scores were therefore compared. For reasons similar to those given above, gender comparisons were omitted, and analyses of covariance by teaching group and citizenship group, with control for Year 1 examination marks, were carried out. Mean change scores were also tested to see if they were significantly different from zero. Full results are in Appendix 2. Summaries are in Tables 4.5 and 4.6

Table 4.5. Comparisons between teaching groups: attitude changes over Year 2.

Variable	Mean differences		Significance of F ratio
	Keller	Non-Keller	
	Chem93	Civ93	
Confidence	08	30##	p < .01
Study methods	.20##	-.07	p < .001
Motivation	.36##	.16##	p < .01
Anxiety	-.23##	13#	p < .001
Control	.34##	.08	p < .05
Subject	41##	32##	ns
N 196	Change significant	# p <.05	## p , .01

Table 4.6. Changes in attitudes: comparisons between citizenship groups.

Variable	Mean differences		Significance of F ratio
	Australians	Overseas	
Confidence	.22	.13	ns
Study methods	.02	.21	p < .05
Motivation	.27	.18	ns
Anxiety	-.03	-.08	ns
Control	20	.21	ns
Subject	.37	.26	ns
N 196			

No interactions were significant.

The change scores for teaching groups show a pattern similar to that found for absolute levels at the end of Year 2. That is, being in a Keller Plan is associated with stronger growth in study methods and motivation scores, but confidence does not grow in this group, and anxiety increases, in marked contrast with positive changes in the traditionally taught group.

The lack of difference between citizenship groups in most areas is not very surprising, but it is of interest that interactions were not significant, which indicates no differential influence of teaching method across groups.

Discussion

The lower confidence and anxiety scores of Keller Plan groups seem to indicate quite clearly that the plan is seen as a considerable challenge by the students. This is not surprising, because the mastery requirement makes a Keller Plan subject far more demanding than a traditionally taught subject, particularly in mathematics, which is quite easily passed by the clever and lazy. The standards of the present implementation were very high, which increases demand. Given the importance of the confidence-anxiety space in studies of total populations, one must take these lower scores quite seriously, as indicating possible disadvantage. On the other hand, the study methods and motivation scores indicate that the Keller Plan group is making a rational and appropriate response to the challenge, taking out anxiety by diligence and a deeper approach. One should note that the overseas students seem also to use anxiety to some profit. The control scores also favour the Keller Plan, indicating greater taking of responsibility for one's studies. The balance between results can only be properly evaluated in conjunction with achievement results, and full discussion is deferred until they have been presented.

The lack of gender differences in attitudes is a negative result of some importance, because it means that the decision to implement a Keller Plan is not complicated by issues involving even a moderate number of different attitudinal effects for males and females. The single significant interactive difference indicated no difference between men's study methods, and some advantage for women's in Keller organisation, which implies no difficulty.

Conclusion

The attitude results support the postulated favourability of the factors assessed. The opposing directions of the group comparisons therefore represent a problem, when considered on their own. For example, the greater confidence of the traditionally taught group is of high potential influence. But when attitudes are placed in the context of the achievement and interview results to be presented in the next chapters, issues are much clarified.

Chapter 5. Achievement: results

Introduction

The chapter covers results involving achievement scores. Results of group comparisons are presented first. Correlational relationships between attitude and the achievement scores specially constructed for the study are then examined.

Comparisons

Achievement scores for the three main Year 2 groups, and for the two groups who went on to Year 3 mathematics, were compared, with the Year 1 examination score as covariate. It should be noted that the Year 3 results involving teaching groups are intrinsically capable of giving a conclusion only if they show significant differences. That is, a finding for or against the group with two years in the Keller Plan would contribute to the argument about the plan's merits, but findings of no differences does not permit an inference because both positive and negative conclusions are compatible with such a result.

Ideally, one would have preferred, here again, a three-way factorial analysis, by teaching group, gender and citizenship, but small cell sizes caused difficulties. Piloting of the three possible choices of two main factors indicated that there were no differences between citizenship groups. It was therefore decided that any important differences would be those found in the two-way analysis of covariance by teaching group and gender, with the lack of differences associated with citizenship being a negative finding in its own right. Results for the three choices of pairs of factors are recorded below.

Comparisons involving citizenship groups

As discussed above, the possibility of differences between citizenship groups was always one that needed consideration, particularly in view of the associations between culture and approach to studying described in the review of literature. It should also be recalled that differences in attitudinal factors were in fact found in the present study. That is, there are plausible reasons for checking whether achievement differences are associated differently with the two teaching methods in the two citizenship groups, because there are cultural and selection differences that distinguish them. What is more, any language difficulties experienced by overseas students could interact with the two different methods of transmitting course material, because written notes are the main source in the Keller Plan, and spoken lectures in the traditional subjects. It is also possible that a break with traditional methods might be found disturbing by the overseas students, particularly because of the face to face immediate assessment in the Keller Plan. On the other hand, the general high level of selection, and the heavy engineering workload could iron out differences in selection and attitudes, and one might also assume surviving overseas students have good social skills. Prediction was therefore difficult. But when results were tested, they showed no significant achievement differences associated with citizenship, either as main factor or in interaction with other factors. Details of the results involving citizenship are given in Tables 5.1 to 5.4. The interest of the negative result is discussed in Chapter 7.

Citizenship and gender

Remarks on gender differences are deferred to the section on teaching group and gender.

Table 5.1 Year 2 achievement: comparisons by gender and citizenship with control for Year 1 examination marks.

A. Means

	Women	Men
Australians	66.28 (116)	58.79 (355)
Overseas	64.45 (33)	61.77 (69)

B. Analysis of covariance

Source of variation	Sum of squares	df	Mefan square	F	Significance
Covariate					
Y1 exam	45441.9	1	45441.9	218.2	.000
Main effects					
Gender	1759.9	1	1759.9	8.5	.000
Citizenship	83.5	1	83.5	.4	.53
2-way interaction					
Gender x Citizenship	406.2	1	406.2	2.0	.16
Residual	118276.6	382	208.2		

Table 5.2 Year 3 achievement: comparisons by gender and citizenship with control for Year 1 examination marks.

A. Means

	Women	Men
Australians	62.28 (22)	63.06 (77)
Overseas	63.71 (14)	68.14 (22)

B. Analysis of covariance

Source of variation	Sum of squares	df	Mean square	F	Significance
Covariate					
Y1 exam	7612.4	1	7612.4	23.42	.000
Main effects					
Gender	295.6	1	295.6	.91	.34
Citizenship	170.5	1	170.5	.53	.47
2-way interaction					
Gender x Citizenship	86.9	1	86.9	.27	.61
Residual	44528.5	137	373.4		

Citizenship and teaching group

Remarks on differences associated with teaching groups are deferred until the section on teaching group and gender.

Table 5.3 Year 2 achievement: comparisons by teaching group and citizenship with control for Year 1 examination marks.

A. Means

	Keller	Non Keller	
	Chem93	Civ93	Chem92
Australians	65.41 (128)	57.21 (225)	61.92 (119)
Overseas	67.09 (35)	57.92 (40)	63.85 (27)

B. Analysis of covariance

Source of variation	Sum of squares	df	Mean square	F	Significance
Covariate					
Y1 exam	45445.1	1	45445.1	227.1	.000
Main effects					
Teaching group	6977.0	2	3488.5	11.9	.000
Citizenship	37.8	1	37.8	.2	.66
2-way interaction					
Teaching group x Citizenship	23.6	2	11.8	.1	.94
Residual	113472.0	567	200.1		

Table 5.4 Year 3 achievement: comparisons by teaching group and citizenship with control for Year 1 examination marks.

A. Means

	Keller in Years 2,3	Keller in Year 3 only
	Chem93	Chem92
Australians	63.53 (51)	62.66 (56)
Overseas	65.00 (22)	68.64 (14)

B. Analysis of covariance

Source of variation	Sum of squares	df	Mean square	F	Significance
Covariate					
Y1 exam	7576.5	1	7576.5	23.2	.000
Main effects					
Teaching group	271.9	1	271.9	.8	.36
Citizenship	140.4	1	140.4	.4	.51
2-way interaction					
Teaching group x Citizenship	182.6	1	182.6	.6	.46
Residual	45054.1	138	326.5		

Comparisons by teaching group and gender

The test results presented immediately below, in Tables 5.5 and 5.6 are, as a result of the negative findings on citizenship, the main source of information about group differences in achievement.

Year 2 results are considered first.

Table 5.5 Year 2 achievement: comparisons by teaching group and gender with control for Year 1 examination marks.

A. Means

	Keller	Non Keller	
	Chem93	Civ93	Chem92
Women	67.26 (58)	64.19 (48)	64.17 (46)
Men	64.12 (108)	58.80 (217)	60.02 (106)

B. Analysis of covariance

Source of variation	Sum of squares	df	Mean square	F	Significance
Covariate					
Y1 exam	47340.4	1	47340.4	230.2	.000
Main effects					
Teaching group	5192.7	2	2290.2	11.1	.000
Gender	788.0	1	788.0	12.6	.05
2-way interaction					
Teaching group x Gender	305.6	2	152.8	.7	48
Residual	11845.4	576	205.5		

The results for Year 2 indicate clearly higher achievement in the Keller Plan group, even with control for Year 1 examination scores. This is therefore the most important finding on achievement, and is consistent with predictions made by supporters of the Keller Plan.

In this year, also, women's scores are significantly higher, and the control is more important here, because their Year 1 examination scores were also higher. That is, without the control, one would only be able to say that it appears that the women were somewhat more highly selected than the men, at least with respect to previous mathematics performance. But with the control, one may say that their higher achievement goes beyond selection effects; that they are coping better with their mathematics even at the same level of selection. The result is, however, open to explanation in terms of selection by other factors. Given the traditionally male image of engineering degrees, it is reasonable to suppose that women's choice of such degrees has been arrived at with rather more deliberation, and adopted with more commitment than men's. A self-selection with respect to commitment and good organisation may therefore be operating, to such an extent as to explain better performance even beyond control. In any case, for the present study, the most interesting finding involving gender is the negative finding of no significant interaction between gender and teaching method.

Year 3 results are added mainly for interest, because they afford only the suggestion of a conclusion, as described below.

Table 5.6 Year 3 achievement: comparisons by teaching group and gender, with control for Year 1 examination marks.

A. Means

	Keller in Years 2,3	Keller in Year 3 only
	Chem93	Chem92
Women	66.00 (24)	58.63 (19)
Men	62.82 (50)	60.02 (50)

B. Analysis of covariance

Source of variation	Sum of squares	df	Mean square	F	Significance
Covariate					
Y1 exam	7642.8	1	7642.8	24.0	.000
Main effects					
Teaching group	176.9	1	176.9	.6	.46
Gender	226.4	1	226	.7	.40
2-way interaction					
Teaching group x Gender	795.4	1	795.4	2.5	.12
Residual	43871.3	138	317.9		

No results here are significant, but the relatively low probability of the F ratio for interaction suggests that there might be some advantage in longer experience of the Keller Plan for women only. On the other hand, the samples of women were quite small, so that chance inclusion of particular extreme achievers could have influenced the result. It follows that the Year 3 results cannot be said to give any definite information. That is, the lack of significant differences could be used for the Keller Plan (effective even in just one year) or against (not more effective over a longer period). It does not follow that the comparison should not have been made, because the other possible outcomes could have given information.

Relationships between attitudes and achievement scores

It was possible to relate attitude scores from the end of Year 2 to achievement scores for that year. Correlations were calculated separately for the three teaching groups. Results are in Table 5.7.

Table 5.7. Correlations, Year 2: Attitude and achievement scores.

	Keller	Non Keller	
	Chem93	Civ93	Chem92
Confidence	53**	.35**	.41**
Study methods	.23**	.06	-.05
Motivation	.28**	.10	.04
Anxiety	.42**	26**	.46**
Control	.22**	.21**	.19
Subject	.22**	24**	33**
Maximum N	128	167	76
	** $p < .01$		

The patterns for confidence, anxiety, control and subject perception are predictable and they are also homogeneous across groups, because the coefficients do not differ significantly. The apparent difference of pattern for motivation and study methods is real to the extent that significance tests (of paired Fisher transforms) indicated that the differences between Chem93 and Chem92 were significant, even though those between Chem93 and Civ93 were not. It follows that there is partial evidence that achievement in the Keller Plan subject was more strongly associated with intrinsic motivation and more diligent and active study methods. Such evidence is reinforced by the attitude changes over Year 2, which showed significantly greater increases in study methods and motivation scores among the Keller group. Since intrinsic motivation indicates a deeper approach, and the study methods scale includes items about studying for real understanding, as well as about diligence and taking action, these differences are clear positive evidence in favour of the Keller Plan, in a way that is predicted by its underlying rationale. It is also of considerable interest to recall Svensson's (1977) early results in the field, which indicated that a deeper approach was a great facilitator of diligence, essentially because studying without the purpose of real understanding is so much more tedious. This is of considerable importance in mathematics learning, because mathematics is so inaccessible to rote learning, and is most effectively understood by using it in problem solving, which has active study as a prerequisite.

Conclusion

The Year 2 achievement results therefore favour the Keller Plan quite clearly. They also indicate that it is associated with certain approaches to study that have been found beneficial to outcomes. There is therefore evidence that provides a considerable counterweight to the association found between the demands of the Keller Plan and lower confidence among students. The interview results provide some further indications that lower confidence is not uniform, and that the students in the plan tend to value it, while those not in the plan strongly criticise the most traditional aspects of their teaching regime. Once the details have been presented in Chapter 6, the full discussion can be concluded.

Chapter 6. Interviews: results

Introduction

The chapter gives an account of the results of the interview substudy. Full details of response patterns are in Appendix 3, and only rough numerical indications are given here. The interviews were carried out mainly to illuminate the results of the attitude questionnaires by asking direct questions about students' experience in the actual Year 2 mathematics subjects they took, and adding questions about changes of approach. They were also intended to allow spontaneous comment. The items deal with attitudinal factors that overlap those studied by questionnaire, and also allow direct questions about experience in, and evaluation of, the specific mathematics subjects studied in Year 2. A few overseas students were included, to allow for the possibility that there were issues important to them that were not covered by the questionnaires.

Findings

Both positive and negative findings are of interest here.

First, it is worth noting that the students in the sample tended to have very similar educational backgrounds, because all of them had Australian university entrance qualifications (q3), and all had done Mathematics 1 at a university in Sydney (q4). They also showed not very dissimilar general attitudes to mathematics (q5), except that the Keller group showed more interest (q5b). Since the information was collected after experience in the two different Year 2 subjects, it is hard to say whether this reflects an initial difference.

Also, both groups tended to show a high level of independence in their mathematics studies, stating that the level required was either good or satisfactory (q9) and not finding it necessary to consult teaching staff much (q13a). If help was obtained, it was mostly from friends (q13b), probably reflecting the high cohesiveness of engineering groups, whose members all do the same programme. Both groups also emphasised (q14b) the importance of good organisation and determination.

It is also worth noting that neither group said their mathematics was too heavy or too hard, but most agreed that the total engineering programme was very heavy, and tended to interfere with their mathematics studies. This means that engineering students find it hard to follow interests, and possibly also to try for deep understanding. Thus a deeper approach is a real achievement in such groups, which makes the findings favouring the main Keller Plan group in motivation and study methods more important.

Differences in study approaches between teaching groups were also indicated by the interview material. The Keller Plan group more often said their interest (q6b) and understanding (q10a) had increased in Year 2, and that they kept up to date better (q6a) and studied both more (q7b) and at a higher level (q12c) than they had in previous mathematics subjects. On the other hand, there was no difference in how well the groups said they remembered their mathematics (q10b), and one Keller student confessed that she tended to cram and then forget (q10a). The non-Keller group found assignments helpful (q12c), indicating that some pressure from in-session assessment was found useful in all cases.

The answers involving confidence did illuminate the quantitative results, because the Keller group showed greater range rather than universal lowering of confidence (q12e). That is, the group was split, with just over half claiming higher confidence, and a third lower. It is reasonably clear that most found the Keller Plan quite challenging, because about half stated that the mastery level required was hard or too high (q12a). On the other hand, in the same context, two thirds of the group said they had found it good for them. The non-Keller group all found their required level of achievement presented no real problems (q12a), but about a third mentioned anxiety about assessment mainly by final examinations (q12e).

The preceding evidence is on the whole not unfavourable to the Keller Plan, but the question about fairness of assessment does indicate continuing difficulties, because just over half of the group expressed doubts about uniformity of marking standards, in contrast with the other group, who were all satisfied that assessment was fair.

The last point is definitely unfavourable to the Keller Plan, even though the students may not be entirely justified, given that the checking of scripts done for the present work did not indicate any clear discrepancies between marking standards. But the last question to be described (q11) seems to afford very important evidence in favour of the Keller Plan. In this question, students were asked how they felt in general about the way their Year 2 mathematics subject had been taught. All but one of the Keller group had found it good, and just over half said it was better than traditional teaching. In the other group, one student had no opinion, and all the rest stated that lectures were not much use to them, though tutorials and problem-solving classes were. Assessment of lectures ranged from preferring fewer lectures (1) or printed notes (1) to finding them too hard (1) or unsatisfactory (1) through to finding them a waste of time (2). It should be noted that the question was very open, and the students' opinions were quite spontaneous.

It should also be noted that the interviewer was not involved in the teaching of either subject.

Conclusion.

The interviews served to qualify results about confidence and anxiety in a way relevant to the final discussion. That is, the effects of teaching methods on confidence showed opposite clusterings that indicated that lower confidence in the Keller Plan was not uniform This point is important, and is taken up in the main discussion. What is more, the anxiety associated with one final examination was mentioned as a disadvantage in the traditionally taught group. The overall evaluations of method were highly illuminating, because the Keller Plan group almost all endorsed the plan, while the traditionally taught group spontaneously and almost uniformly stated that they did not find much profit in lecture classes.

Chapter 7. Discussion

Introduction

The main discussion in the chapter is concerned with the meaning of the undoubtedly higher achievement scores for Keller Plan students. The validity of the scores is reviewed, as are the precautions taken to avoid sample bias. The achievement results are then placed in the context of the attitudinal results, and a final balance in favour of the Keller Plan is presented. Comments are also made on relativities between groups defined by gender and citizenship.

Achievement

Teaching group comparisons

Year 2

Results for achievement are fairly clear, as long as one accepts the validity of the comparison method. A summary of the arguments used to justify it, while simultaneously acknowledging limitations, is given below.

The scores

No further elaboration of the method used to construct the achievement scores needs to be given here, but a resume of the justificatory arguments for validity of scores seems appropriate.

The main issue is the validity of using scores on different tasks in comparisons, an inevitable problem if one cannot obtain a final common test. Administrative difficulties apart, there is one compelling reason for not making the attempt to organise a common test when attempting to evaluate a Keller Plan implementation. The point is that the different teaching methods used would always mean that a common test would be one specially set and not involving credit for the degree, at least for the Keller group, and possibly also for controls. As a result, all or some of the sample would be in an artificial situation, where motivational influences would inevitably be different from those acting on students being assessed for credit. Artificial common tests may be unbiased over samples if they are artificial for all groups being compared, but even then one cannot say that using them allows one to evaluate the real demands of subjects taught by different methods. Since real demands in real implementations are the important object of study, the whole purpose of the comparison would not be adequately served.

The intentions of the present study therefore force the use of material generated in the process of real assessment. What is more, the best control group was obviously the group in the same subject from the year prior to the initiation of Keller Plan teaching. The use of answers to different sets of mathematics questions is forced as a result. It is the validity of the scoring method based on broad levels of coherence in students' answers that is in question. The essential feature of the original argument given above in its favour was that limitation of variability in subject matter and testing level makes validation easier, and that the resulting restriction of the difficulty of the validation task makes it possible to use methods that are very similar to those used in normal careful marking of examinations. Thus it is first assumed that validity of marking of individual questions by professional mathematicians can be justified. This is taken as largely true, because answers to mathematical tasks are highly convergent, so that both reliability and valid assessment of quality are easier to achieve than in more divergent subjects. What is more, validation is confirmed by cross correspondences and links with performances on related tasks.

Validation of a rough ordering of levels of quality across different, but very similar, tasks, is therefore the next problem. It is claimed here that it can be achieved, again via convergence and via methods that resemble reliability testing within a single normal marking task. That is, it is justified by the details of case law applying to an exhaustive set of variant answers. No claim for total exhaustion of all possibilities is being made here. Instead, exhaustion is of the finite number of variations found, and the claim is that the case law gives enough detail to allow consistent scoring of any new response. Clearly, here, there is a further assumption, that responses will fall into quite recognisable patterns. It has been pointed out above that this is indeed so for the finite universes involved, and that some decades of marking experience support the assumption. That is, first, normal marking indicates that students' errors in answering a single question are usually not very original, once one has seen about fifty scripts, and that patterns for other questions examining similar material are generally similar. And second, patterns were in fact found to be both restricted and of very similar types in the material assessed for the present project.

Validation is therefore claimed to be at least as good as that required for the comparison, for example, of examination scores from adjacent years when syllabuses have not changed, and the ability and preparation of the student groups can be assumed not to have changed significantly. With large groups of engineering students in reasonably stable degree courses, the latter requirement is highly likely to be satisfied.

The achievement scores are therefore, it is claimed, of a validity comparable with that of scores normally used without much question in educational research, so that one may proceed to examine results involving them.

Inclusion precautions

Tested material

All achievement scores were calculated over the full amount of material set for the subject. That is, what was compared was the average level of mastery over all topics in the syllabus. That means that those Keller Plan students who did not do a test on a given topic received a score of zero for that topic, which then contributed a zero to their average score. The intention here was to make sampling of total achievement tasks fair, so that an appropriate comparison might be made between the Keller group and the traditionally taught groups, whose final examinations tested their whole syllabuses. The scores therefore, in both cases, represent levels of achievement over the full body of material set, under formal examination conditions.

Inclusive samples.

A second precaution was taken, to cope with the problem of possibly different patterns of subject discontinuation among students taught in different teaching plans. Since test results are immediately available to Keller Plan students, they have much earlier feedback about how they are doing, than do traditionally taught students. If they get very far behind the schedule implied by the normal running period of the subject, this is a very clear signal that their prospects are bad, and they may choose to discontinue the subject more readily than students who are just as far behind in a traditionally taught subject, but are not getting external signals that make this apparent. What is more, students who discontinue early enough may do so without a failure in the subject being recorded. It is therefore possible that, if such cases were excluded, the Keller plan sample might be biased upward by losing more poor performers than control groups. In the present study, great care was taken to include all those Keller plan students whose presence in the subject enrolment group could be detected by their having attempted at least one test. All topics on which a test was not completed received a zero score, as described above, representing failure to learn that part of the material. Since no similar information was available for traditionally taught groups until the mid year examinations, one may assume that any bias that may have occurred as a result of not counting discontinuing students would favour the traditionally taught groups.

Group selection

Different types of engineering degrees may have different levels of selectivity, depending on where the brightest students apply most often for admission. It is recalled here that all comparisons were controlled for Year 1 mathematics examination scores in two ways. First, the examination scores of the teaching groups were compared, and no significant differences were found. Second, all later comparisons used the Year 1 examination score as a covariate. A third fact that should be recalled is that one of the non-Keller groups came from the previous year's intake to the same range of degree courses as the Keller Plan group. Gross distortion of results due to intake selection may therefore be assumed not to have occurred.

Conclusion

The above summaries indicate that achievement scores were probably of validity similar to that accepted for a large proportion of research studies in education. What is more, the precautions taken to control sampling of achievement tasks and group selection seem to cover the obvious issues. One may therefore have more confidence in the findings than would otherwise have been justified.

The achievement results

The comparisons of second-year groups clearly favour the students taught by the Keller Plan. That is, using the criteria chosen, the Keller group shows greater mastery of the subject matter of the whole subject, demonstrated under traditional test conditions. The final performance is what is counted, and it is clear that superior performance is likely to reflect the more stringent requirements of the Keller plan. Whatever the mechanism, the Keller plan seems to have got more out of the students.

The correlations between the achievement score and the attitude scores may also be taken as favouring the Year 2 Keller Plan, because success in the subject appeared to require deeper approaches to study. That is, the achievement score was, in contrast with the other groups, significantly associated with intrinsic motivation, and diligent study methods, which included a determination to achieve real understanding of the work. The significance of the differences between correlations for the two Chemical Engineering groups confirms the conclusion in favour of Keller Plan teaching. This conclusion is of some importance, because, as noted in the review of literature, interest and intention to pursue understanding, as part of the deep approach to study, are predictably part of the process of achieving high quality of outcome. What is more, the literature tends to justify the conventional wisdom about diligent study methods, if methods are assessed in suitable generality, emphasising activity and good organisation rather than specific methods. The clearest results, however, have most often been in the humanities, arguably because the commonly used instruments are more suited to the area than to science and mathematics. For example, one may recall that Watkins (1982) did not find evidence even for the existence of a deep approach among Science students. It is therefore of some interest that the present study, with instruments adapted to mathematics, affords results that can be interpreted as showing that a deep approach is associated with the mastery requirement of the Keller Plan.

Year 3

The ancillary material comparing the third-year groups, Chem92 and Chem93, both of whose third year mathematics was taught in a Keller plan, indicated no significant differences between group achievement scores. Thus one cannot say that Chem93, who had had Keller organisation in their second year as well, showed any significant long term advantage, even though one might have postulated some effect, either from practice in the method or through better retention of the second year subject matter. One should note that the third year groups are even more highly selected than the second year groups, and that the second year subject is not very difficult mathematics as second year university mathematics goes. It is possible that this implies a ceiling on the effects that second year teaching organisation might have. That is, the stronger students might retain second year work quite easily under either method. Whatever the explanation, however, there was no evidence to suggest a significant effect of the second year Keller organisation on third year performance in a similarly organised course. The result is essentially a piece of negative evidence, because the lack of difference is open to different interpretations, rather than supplying material for an argument in favour of one conclusion. That is, one should recall that it could be interpreted as showing that Keller Plan teaching has no long term effects, but it could also be interpreted as indicating that the plan has effects favourable enough to mask any effect due to longer experience of the teaching method. Because there was no direct testing, the present results cannot contribute to the question of retention. There is therefore only one safe conclusion to be drawn: the results do not give any definite information against the Keller Plan. Thus the absence of differences neither supports nor clearly contradicts the findings of Watson (1986), who obtained lower scores for Keller Plan students on a retention test taken eight months after the completion of the respective courses.

Comparisons between women and men

In Year 2, women's achievement results were significantly better than men's, even with control for first year mathematics performance. Given the control, it is not obvious that one should attribute the result to selection of women of higher mathematical ability, even though women's higher Year 1 examination marks mean that this is a possibility. On the other hand one cannot ignore the fact that engineering degrees are less traditional for women than for men, so that stronger commitment is required of women who apply for such degrees, which would be quite sufficient to explain the result. Since the F ratio for interaction between teaching group and gender was far from significant, the contribution of the results to the present study is a negative finding. Thus one may conclude that, in Year 2, the women engineering students tended to achieve better than men at a similar level of initial achievement, but this appears to be independent of the teaching methods examined. The most probable explanation, that of selection by commitment, makes the finding context dependent.

In Year 3, gender differences had disappeared, but the women's samples were smaller, so that results could have been affected by chance inclusion of extreme achievers, high or low. A similar argument implies that one should not make too much of the relatively low probability of the F ratio associated with interaction between gender and group, although it does point towards a possible greater benefit of longer Keller Plan teaching for women.

Comparisons between citizenship groups

The lack of significant differences between the Australian group and the overseas group is actually of some interest, because, as pointed out above, there are selection differences between groups that are predictable from the overseas' students' circumstances, and verified by comparisons of initial data collected for the present study. First, it has been pointed out that the overseas students have very strong reasons for commitment and a culture that gives a high value to study, and are implicitly highly selected for organisational and life skills. Material from the first data collection did indicate that they were at an initial advantage with respect to both previous achievement and a range of attitudinal factors. Second, it was noted above (Chapter 2) that one could not assume Keller Plan teaching would have similar effects on different cultural groups. In view of both considerations, the negative finding is of some interest.

Conclusion

It seems clear that the achievement results favour the use of the Keller Plan in the Year 2 subject. It is also clear that the ancillary Year 3 results are a purely negative finding, and do not afford an argument, in either direction, about long term effects.

The discussion above included attitudinal material only as far as correlation differences were concerned, and the pattern also favoured the Keller Plan. But a complete examination of attitude results remains to be made. The main point is to find out if there are any attitude differences that might offset the achievement advantages found for the Keller Plan.

Attitudes

Meaning of scores

The favourability of high scores on the attitude scales as an expression of straightforward job satisfaction is clear from the content of the items.

The significant positive correlations of the scale scores with examination marks indicate that they are positively associated with achievement in mathematics. It should be noted that the present sample, consisting of engineering students who have gone on to a second year of mathematics, is clearly somewhat upwardly selected with respect to the overall Year 1 group, so that the size of correlations is likely to have been reduced because of reduced variability in their Year 1 examination scores. Hence there is no problem about accepting the attitudinal factors as favourable, even though it is freely admitted that no conclusion of causality can be drawn.

Group comparisons

Initial attitude scores

It is recalled that there were no significant attitudinal differences between the Keller group Chem93 and the non-Keller group Civ93 at the beginning of their Year 2 mathematics work. Hence it is more justifiable than it would be without the check, to claim that any subsequent group differences are related to the teaching method encountered in Year 2.

On the other hand, initial comparisons between citizenship groups did show differences, favouring the overseas students in confidence, study methods, motivation and willingness to take responsibility for their own studies. Such initial differences are highly compatible with what one would predict from background knowledge about the position of the overseas students, so that similarly significant absolute differences at the end of the year are predictable, and it is likely that results involving interactions with teaching group, or changes in attitudes are those mainly of interest.

A preliminary note on gender

There was only one significant attitudinal difference associated with gender. That was a significant interaction between study methods and gender, with the Keller Plan apparently more favourable to women than to men. As an overall general finding of minimal differences, the pattern is of some interest, because of the wide extent of gender differences in attitudes found in more broadly based samples. It is, however, interpretable as merely a confirmation of the underlying group selectivity. It follows that, as far as attitudes are concerned, the main comparison results in the present study may be taken to be those not involving gender. It was explained above why this led to the presentation of attitudinal results analysed without introducing gender as a factor.

Teaching groups

One can now evaluate the attitudinal differences of Year 2. The frequency of significant changes in attitudes indicates that one may safely say that students' experience in Year 2 had a significant influence on their attitudes. Given that changes over the year were very different in different teaching groups, it is reasonable to relate the changes to teaching.

It seems suitable to look first at those attitudes that are components of students' approaches to study. The problem that has to be dealt with is what conclusion to draw from the overall set of results, given that significant differences from different areas have opposite directions. That is, to what extent can one conclude that the higher scores of the Keller Plan group in motivation, study methods and control, which it seems justified to regard as advantageous, have their importance reduced by the group's lower scores in confidence and anxiety? The point here is that confidence is the dominant factor, in approaches to study, that emerges in the present work, and has also been established as supremely important in studies of secondary mathematics learning conducted on samples representative of whole populations. The now classic studies in this area are the Fennema-Sherman studies of the later 1970s (Fennema & Sherman, 1977, 1978) and the very large scale, very carefully sampled, confirming work done by Armstrong (1979) as part of the second American National Assessment of Educational Progress. In this work, confidence was a most important correlate of achievement, and group differences indicated that levels of confidence were to some degree independent of previous achievement, which admits the possibility of confidence as an input to achievement rather than just a straightforward reflection of it. Fennema's (1985) model proposes confidence as the substrate of autonomy in learning, which is of vital importance because the most effective learning procedure is problem solving, which is necessarily autonomous. The large scale American studies did not treat anxiety separately, because their instruments showed scores almost completely in a straight line relationship with confidence scores. Hence, in these studies, anxiety results are merely mirror images of confidence results, and hence anxiety is a disadvantage. Further, in work on tertiary students' approaches to study, starting with the close analyses of Fransson (1977), and developing in the work of Ramsden and Entwistle (1981), and Biggs (1982), the results indicate that anxiety is associated with surface approaches and reduced effectiveness. The great importance of the approach used in this work, however, lies in its emphasis on considering whole configurations of factors rather than single factors.

If, in the present work, one looks at the total context, there is a perfectly straightforward, rationally accessible environmental factor influencing attitudes. Quite simply, the Keller Plan mathematics subjects are far more demanding than the traditionally taught ones, because the mastery requirement is at a high level and enforced quite strictly. The sort of error that usually led to rejection of a Keller Plan student's test work was very frequently not bad enough by normal marking standards to impose a failing score in a traditional examination. Somewhat lower assurance about ease of coping with Keller mathematics is therefore quite rational. It should also be recalled that when the non-Keller groups were combined, the confidence difference lost significance.

One should also note that the confidence and anxiety scores of the Keller group are by no means low in absolute. In this connection, it is quite illuminating to examine distributions of scores. For example, the distribution of confidences score appears below in Figure 7.1.

One should first recall that the mean score is 3.34, above the neutral value of 3. Second, one should note a certain spread in both directions, as if reactions were somewhat polarised, with plateaux indicating upward and downward shifts each side of the peak. Interview results should be recalled here, because there was a clear split in the interviews with just over half stating their confidence had increased, but a substantial minority stating it had decreased. Hence the reaction is not uniform, and the average level far from depressed. It appears, therefore, that the confidence of the Keller group was not lowered to any extreme level. Similar remarks apply to anxiety scores, even though the mean was slightly below the neutral point. It should also be noted that the present study is concerned with a total group already very highly selected for confidence about, and achievement in, mathematics and science, so that it is hard to say how predictive one should expect results for representative populations at secondary level to be.

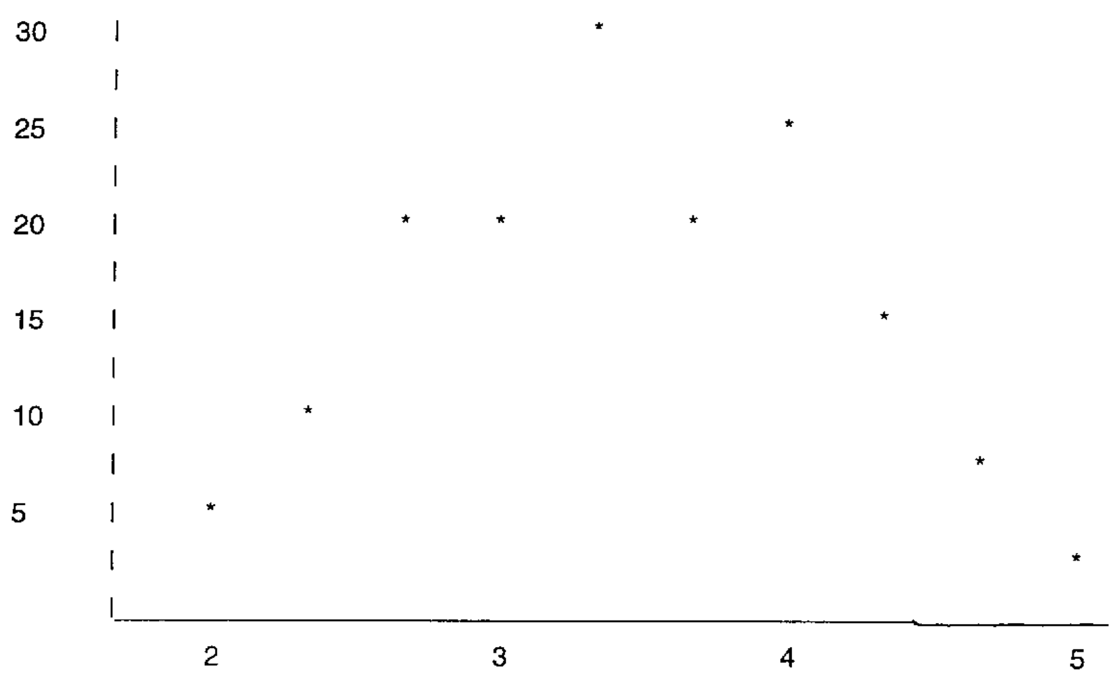


Figure 7.1 Confidence score distribution: Chem93

The next thing to be noted is that lowered confidence scores are accompanied by higher study methods scores, and a greater feeling of responsibility for one's studies. This indicates an effective coping response to the greater challenge of the Keller subject. Confirmatory details of greater effective coping can be found in the interview material. One should also recall that the group of overseas students had higher scores for confidence and study methods, but lower anxiety scores. The anxiety scores in this case accompany a theoretical and practical allegiance to the conventional wisdom about hard work, which leads one to believe that anxiety which is increased, but not to excess, and worked out by diligence, is no disadvantage. And the demands of the subject appear to have made the intrinsically motivated deep approach more closely related to achievement in the Keller group.

It seems that one may therefore conclude that, if the right configuration of requirements and complementary responses is present, greater anxiety and lower confidence, within limits, need not be a disadvantage. As always in work about students' approaches to study, it is the total configuration of the approach components that counts. And finally, effects on the Keller Plan group seem to be justified by their higher achievement scores.

The above conclusion, however, would be much weakened if the lower confidence and anxiety scores made such a difference to Keller students' self-perceived job satisfaction as to result in observable misery. Observables here can be sought in the subject evaluation scores and the interview material.

The subject evaluation scores in fact do not discriminate between teaching groups. Supporters of the Keller plan might be somewhat disappointed in the present result. But the actual mean score of the Keller group was almost identical with the higher of the other two group means, and all are somewhat above the neutral point of the scale. One can therefore conclude that there is a universally mild positive evaluation of the subjects under consideration, at about the same level across subjects.

Is one then to conclude that students feel only that the Keller Plan is no worse than usual, even though usual is not seen as very terrible? Here the interview results make some contribution, because only one Keller Plan student failed to show some enthusiasm for the teaching method. More than half the students saw the Keller Plan as better than lectures, and all but one of the rest showed satisfaction, though in some cases qualified by perceiving it as a greater challenge. But positive results from the Keller group are actually not as interesting as the results for the corresponding question given to the group that was taught in traditional lecture-tutorial format. Despite the mildly positive overall questionnaire scores, when the question of the organisation of teaching in lectures and tutorials was raised explicitly in the interviews, the overwhelming consensus in the (admittedly small) interview group was that lectures were not considered a very useful part of the teaching, in marked contrast to classes devoted to problem solving. Given also that the interview material confirmed the greater diligence of the Keller group, one must conclude that the evidence suggests they felt they were using the time set aside for the subject rather more profitably than did the traditionally taught group.

The lack of difference in the questionnaire scores, in conjunction with the interview material, therefore allows a conclusion mildly in favour of the Keller Plan on subject evaluation. Given this, one may also conclude that the group's lower confidence and anxiety scores do not indicate obvious overall misery ascribable to the plan's organisation of teaching.

Comparisons between citizenship groups

One should first note that there are almost no differences between citizenship groups in the attitude change scores, the only significant one being in study methods. What is more, changes showed no interaction with teaching group. One may therefore conclude that attitudinal differences are relatively independent of the type of Year 2 mathematics teaching experienced by students.

Turning now to differences at the end of Year 2, one finds the overseas group means higher for study methods and motivation scores, but lower for anxiety scores. All of these relativities were present in initial scores, significantly for the first two, and no interactions with teaching group were significant. The results therefore do not seem to indicate any different effects of teaching methods on the two main citizenship groups. It is worth noting that the only overseas student who mentioned the language factor in the interviews actually found the Keller written material hard.

The actual differences found are as one would predict from background knowledge, both of the cultures from which the overseas students come, and of the pressures implied by their having come to study as full fee paying students in Australia. The point of interest of the pattern of differences is the combination of greater anxiety and greater diligence, which is similar to that found overall for the Keller group. This pattern among the overseas students also supports the argument that the challenge of high demands can be appropriately met in such a way that anxiety may be turned to good account.

Limitations of the study.

A first limitation of the study is one common to many educational investigations, that is, that naturally occurring groups have had to be used, so that assignment to treatment or control groups cannot be claimed as random. This limitation is not considered serious.

Second, the sample of chemical engineering students is quite high in ability, even among university students. On the other hand, the main investigation dealt with Year 2 students, not all in chemical engineering, and not unusually highly selected for their year position. The relatively higher selection of women and overseas students was allowed for in the design. Nevertheless, it does not follow that groups who go on to a second year of university mathematics are entirely comparable to a first year group, or a group of significantly lower mathematical ability. Groups of very high mathematical ability are also not necessarily comparable.

Further research

Given the limitations noted above, one obvious choice for further investigation would be the generality of results for groups in earlier stages of university mathematics, or of mathematical ability very different from that of the samples considered in the present work.

In this work, it appears from the above discussion that the Keller Plan implementation studied may be favourably evaluated. That is, the achievement results are clearly favourable, and the attitudinal results on balance also so. The point about meeting challenge with an appropriate response is conditional on the level of challenge being not too extreme. That is, the effects of challenge are probably curvilinear, in a similar way to results from studies of levels of arousal optimal for different tasks. The advantage, in this connection, of the Keller Plan provision for multiple test taking is also clear, because it offers an immediate plan of action for remedying lack of success, and a framework for getting help. Optimisation of challenge and redemption is a field that obviously suggests itself for further research.

A further aspect of optimisation which is brought to attention by the present work, implicitly in its general target, but explicitly formulated in the interviews, is that of the optimisation of the use of time. The main question here is the use of lectures for exposition of material, which the traditionally taught group tended not to find very useful. Given that mathematics lectures tend to be used to convey well defined information, which is quite similar to what could be conveyed by notes or textbooks, it would be possible to replace lectures by printed information and problem solving classes, in a variety of teaching plans. On the other hand, Keller Plan implementations are likely to be the most common naturally occurring examples, so that variety may not be easy to achieve.

Conclusion

The results of the study afford a conclusion in favour of the use of the Keller Plan in tertiary mathematics teaching, subject to the qualifications made at each stage of the text.

The aims of the work, stated at the beginning of Chapter 1, were to investigate the quality of the observable outcomes of different methods of university mathematics teaching, subject to no markedly disadvantageous effect on attitudes. The results clearly favour observable outcomes of Keller Plan teaching, and the discussion above implies that effects on attitudes are not, on the whole, unfavouable.

The consideration of naturally occurring subgroups was more an aspect of control than a central aim. The results emerging were easily explained. That is, overseas students were initially more committed to their mathematics studies, and rather better prepared for Year 1, but not differently affected by Keller Plan teaching. Also, the women in the group tended to show higher achievement, consistently with what one might expect in engineering degrees.

It is of particular interest that the implementation studied was a servicing subject for engineering students, for whom mathematics was a compulsory subject. Because of this, one may take the findings as relevant to a normal world of ongoing teaching, in which groups are not specially selected for interest in mathematics.

Reference list

Armstrong, J.M. (1979). A national assessment of achievement and participation of women in mathematics. Resources in Education ED 187 562

Bar-Tal, D. (1978). Attributional analysis of achievement-related behaviour. Review of Educational Research, 48, 259-271.

Barrett, E.M. & Prokhovnik, S.J. (1980). Reflections on a Keller Plan experiment: Its essence and adaptability. Journal of Personalised Instruction, 4, 100-102.

Bauersfeld, H. (1991). The structuring of the structures. Bielefeld: Bielefeld University.

Biggs, J.B. (1976). Dimensions of study behaviour: Another look at ATI. British Journal of Educational Psychology, 46, 68-80.

Biggs, J.B. (1979a). Individual differences in study processes and the quality of learning outcome. Higher Education, 8, 381-394.

Biggs, J.B. (1979b). The Study Process Questionnaire. Newcastle NSW: University of Newcastle.

Biggs, J.B. (1982). Student motivation and study strategies in university and college of advanced education populations. Higher Education Research and Development, 1, 33-55.

Biggs, J.B. & Collis, K.F. (1982). Evaluating the quality of learning: The SOLO taxonomy. New York: Academic Press.

Boden, M.A. (1974). Piaget. Brighton: Harvester Press

Boud, D.J., Bridge, W.A. & Willoughby, L. (1975). PSI now: A review of progress and problems. British Journal of Educational Technology, 6, 15-34.

Bridge, W. (1976). Self-study courses in undergraduate science teaching. Higher Education, 5, 211-224.

Brook, R.J. (1977). An experiment in self-paced study. Resources in Education, ED 142 242.

Calhoun, J.F. (1973). Elemental analysis of the Keller method of instruction. Resources in Education, ED 088 382

Casanova, J., Casanova, C., Fernandez, L. & Villamanan, M.A. (1979). Teaching thermodynamics to students using the Keller Plan. European Journal of Science Education, 1, 65-69.

Chick, H.L., Watson, J.M. & Collis, K.F. (1988). Using the Solo taxonomy for error analysis in mathematics. Research in Mathematics Education in Australia, May 1988, 34-47.

Cobern, W. (1993). Contextual constructivism: The impact of culture on the learning and teaching of science. In K. Tobin (Ed.) (1993) The practice of constructivism in science education, (pp. 51-69). Hillsdale NJ: Erlbaum.

Coleman, J. S., Campbell, E.Q., Hobson, C.J., McPartland, J., Mood, A.M., Weinfeld, F.D. & York, R.L. (1966). Equality of educational opportunity. Washington DC: United States Government Printing Office.

Collard, G.L. (1989). Lecture versus modified Personalized System of Instruction: Differences in student achievement, attrition, attitudes and other selected variables. Dissertation Abstracts, 49, 1681A.

Collis, K.F. (1972). Concrete to abstract - a new viewpoint. The Australian Mathematics Teacher, 28, 113-118.

Collis, K.F. (1975a). A study of concrete and formal operations in school mathematics: A Piagetian viewpoint. Melbourne: Australian Council for Educational Research.

Collis, K.F. (1975b). The development of formal reasoning. Newcastle, NSW: University of Newcastle.

Condo, P. (1974). The analysis and evaluation of a self-paced course in calculus. Paper presented at the National Conference on Personalized Instruction in Higher Education, Washington DC.

Dahlgren, L. (1984). Outcomes of learning. In F. Marton, D. Hounsell & N.J. Entwistle (Eds). The experience of learning. Edinburgh: Scottish Academic Press, pp. 19-35.

Edward, J. (1976). A comparison of the traditional method and a personalized system of instruction in college mathematics. Dissertation Abstracts, 37, 904A-905A.

Entwistle, N.J. (1992). Student learning and study strategies. In B.R. Clark & G. Neave (Eds.). The Encyclopedia of Higher Education, (pp. 1730-1740). Oxford: Pergamon Press.

Entwistle, N.J., Hanley, M. & Hounsell, D. (1979). Identifying distinctive approaches to studying. Higher Education, 8, 365-380.

Entwistle, N. J. & Ramsden, P. (1983). Understanding student learning. London: Croom Helm

Entwistle, N.J. & Waterston, S. (1988). Approaches to learning and levels of processing in university students. British Journal of Educational Psychology, 58, 258-265.

Entwistle, N.J. & Wilson, J.D. (1977). Degrees of excellence: The academic achievement game. London: Hodder and Stoughton.

Fennema, E. (1985). Explaining sex-related differences in mathematics: Theoretical models. Educational Studies in Mathematics, 16, 303-321.

Fennema, E & Sherman, J. (1977). Sex-related differences in mathematics achievement, spatial visualization and affective factors. American Educational Research Journal, 14, 51-71.

Fennema, E & Sherman, J. (1977). Sex-related differences in mathematics achievement and related factors: A further study. Journal for Research in Mathematics Education, 9, 189-203.

Fennema, E. , Wolleat, P.L. & Pedro, J.D. (1979). Mathematics attribution scale. Journal Supplement Abstract Service, 9, ms 1837.

Flammer, G.H. (1971). Learning as the constant and time as the variable. Engineering Education, 61, 511-514.

Fransson, A. (1977). On qualitative difference in learning IV: Effects of motivation and anxiety on process and outcome. British Journal of Educational Psychology, 47, 244-257.

Friedman, C.P., Hirschi, S., Parlett, M. & Taylor, E.F. (1976). The rise and fall of PSI in physics at MIT. American Journal of Physics, 44, 204-211.

Friedman, C.P., Kaplan, M.N. & Cheatham, C.M. (1979). A comparison of personalized instruction and lecture/discussion courses in electrical engineering with self-selection of treatment. Resources in Education, ED 174 131

Goldschmid, B. & Goldschmid, M.L. (1973). Individualising instruction in higher education: A review. Higher Education, 3, 1-24.

Gorsuch, R.L. (1983). Factor analysis. Hillsdale NJ: Lawrence Erlbaum.

Green, B.A. (1971). Physics teaching by the Keller plan. American Journal of Physics, 39, 764-775.

Harris, M.B. & Liguori, R.A. (1974). Some effects of a personalized system of instruction in teaching college mathematics. Journal of Educational Research, 68, 62-66.

Hassett, M.J. (1978). Effects of student attitude data on the development of a PSI program in mathematics. Resources in Education ED 162 855

Herndon, J. (1974). The way it spozed to be. London: Pitman

Horak, V.M. (1981). A meta-analysis of research findings on individualized instruction in mathematics. Journal of Educational Research, 74, 249-253.

Howarth, M.J. & Smith, B.J. (1980). Attempts to identify and remedy the mathematical deficiencies of engineering undergraduate entrants at Plymouth Polytechnic. International Journal of Mathematics Education in Science and Technology, 11, 377-383.

Imrie, B.W., Blithe, T.M. & Johnston, L.C. (1980). A review of Keller principles with reference to mathematics courses in Australasia. British Journal of Educational Technology, 11, 105-121.

Karp, H.J. (1983). The use of Keller's personalized system of instruction. Resources in Education ED 231 295.

Keller, F.S. (1968). "Goodbye, teacher...". Journal of Applied Behavior Analysis, 1, 79-89.

Keller, F.S. & Sherman, J.G. (1974). The PSI handbook: Essays on personalized instruction. Menlo Park CA: Benjamin

Kellogg, R.J. (1973). Locus of control and teachers' attitudes to student autonomy. Dissertation Abstracts, 34, 2932A.

Kerrigan, J.J. (1976). A comparison of teacher paced and individualized student paced instruction in college freshman mathematics. Dissertation Abstracts, 37, 2038A-2039A.

Klopfenstein, K.F. (1977). The Personalized System of Instruction in introductory calculus. American Mathematical Monthly, 84, 120-124.

Kontogianes, J.T. (1974). The effects on achievement, retention and attitude of an individualized instructional program in mathematics for prospective elementary school teachers. Dissertation Abstracts, 34, 5802A.

Kulik, C.C. & Kulik, J.A. (1976). PSI and the mastery model. In B.A. Green (Ed.) (1976), Personalized instruction in higher education. Washington DC: Center for Personalized Instruction, pp. 155-159.

Kulik, C.C. & Kulik, J.A. & Bangert-Drowns, R.L. (1990). Effects of mastery learning programs: A meta-analysis. Review of Educational Research, 60, 265-299.

Kulik, J.A ,Kulik, C.C. & Carmichael, K. (1974). The Keller Plan in science teaching. Science, 183, 379-383.

Kulik, J.A ,Kulik, C.C. & Cohen, P.A. (1979). A meta-analysis of outcome studies of Keller's Personalized System of Instruction. American Psychologist, 34, 307-318.

Kulik, J.A ,Kulik, C.C. & Smith, B.B. (1976). Research on the Personalized System of Instruction. Programmed Learning and Educational Technology, 13, 23-30.

Lamal, P.A. (1984). Interest in PSI across sixteen years. Teaching of Psychology, 11, 237-238.

Lloyd, M.E. & Lloyd, K.E. (1986). Has lightening struck twice? Use of PSI in college classrooms. Teaching of Psychology, 13, 149-151.

Locksley, N. (1977). The Personalized System of Instruction (PSI) in a university mathematics class. Dissertation Abstracts, 37, 4194A-4195A.

Lu, M.C. (1976). The retention of material learned by PSI in a mathematics course. In B.A Green (Ed.) (1976), Personalized instruction in higher education, (pp. 151-154.). Washington DC: Center for Personalized Instruction.

Malec, M.A. (1975). PSI: A brief report and reply to Francis. Teaching Sociology, 2, 212-217.

Martinez, J.G.R. & Martinez, N.C. (1988). "Hello, teacher": An argument for reemphasizing the teacher's role in PSI and mastery learning. American Journal of Education, 97, 18-33.

Marton, F. & Säljö, R (1976a) On qualitative differences in learning I: Outcome and process. British Journal of Educational Psychology, 46, 4-11.

Marton, F. & Säljö, R (1976b) On qualitative differences in learning II: Outcome as a function of the learner's conception of the task. British Journal of Educational Psychology, 46, 115-127.

Meyer, J.H.F. & Parsons, P. (1989). Approaches to studying and course perceptions using the Lancaster inventory: A comparative study. Studies in Higher Education, 14, 137-155.

Nord, S.B. (1975). Comparative achievement and attitude in individualized and class instructional settings. Dissertation Abstracts, 35, 5129A.

Norusis, M.J. (1990). SPSS base system user's guide. Chicago: SPSS Inc.

Nunnally, J.C. (1967). Psychometric theory. New York: McGraw-Hill.

Pascarella, E.T. (1977a) Student motivation as a differential predictor of course outcomes in Personalized System of Instruction and conventional instructional methods. Journal of Educational Research, 71,21-26.

Pascarella, E.T. (1977b). Interaction of motivation, mathematics preparation and instructional method in a PSI and conventionally taught calculus course. AV Communication Review, 25, 25-41.

Pascarella, E.T. (1978). Interactive effects of prior mathematics preparation and level of instructional support in college calculus. American Educational Research Journal, 15, 275-285.

Pask, G. (1976). Styles and strategies of learning. British Journal of Educational Psychology, 46, 128-148.

Pearson, M. (1983). Approaches to individualising instruction: A review. Higher Education Research and Development, 2, 155-181.

Pegg, J.E. (1985). How children learn geometry: The Van Hiele theory. The Australian Mathematics Teacher, 14, 5-8.

Pegg, J.E. (1991). Learning cycles in mathematics. Seminar given at the University of Western Sydney.

Pegg, J.E. & Davey, G. (1989). Clarifying level descriptors for children's understanding of some basic 2-D geometrical shapes. Mathematics Education Research Journal, 1, 16-27.

Peluso, A. & Baranchik, A.J. (1977). Self-paced mathematics instruction: A statistical comparison with traditional teaching. American Mathematical Monthly, 84, 124-129.

Piaget, J. (1974). The origins of intelligence in children. New York: International Universities Press.

Powell, T.H. (1980). An instructional alternative for college instructors: PSI. Resources in Education, ED 187 284.

Prokhovnik, S.J. (1975). The Keller Plan at the University of New South Wales. The Australian Mathematical Society Gazette, 2, 13-15.

Rainey, V.P. (1980). An assessment of selected measures of success in an individualized mastery learning approach to beginning algebra as a predictor of success in intermediate algebra. Dissertation Abstracts, 40, 3850A-3851A.

Ramsden, P. & Entwistle, N.J. (1981). Effects of academic departments on students' approaches to studying. British Journal of Educational Psychology, 51, 368-383.

Richardson, J.T.E. (1990). Reliability and replicability of the Approaches to Study Questionnaire. Studies in Higher Education, 15,155-168.

Roberson, J.A. & Crowe, C.T. (1975). Is self-paced instruction really worth it? Engineering Education, 65, 761-764.

Roberts, F.C. (1975). A cost-effective analysis and follow up study on a multi-level mathematics instruction system at Antelope Valley College. Resources in Education ED 125 689.

Romiszowski, A.J. (1979). What's happening to individualised mathematics? Programmed Learning and Educational Technology, 16, 146-150.

Roth, C.H. (1973). Continuing effectiveness of personalised self-paced instruction in digital systems engineering. Engineering Education, 63, 447-450.

Roueche, J.E. & Mink, O.G. (1975). Toward personhood development in the community college. Resources in Education, ED 114 137

Roueche, J.E., Mink, O.G. & Abbott, M.L. (1975). Effects of individualized instruction on control expectancy: A field test. Resources in Education, ED 114 138.

Schielak, V.P. (1983). A personalized system of instruction versus a conventional method in a mathematics course for elementary education majors. Dissertation Abstracts, 43, 2267A.

Semb, G.B. (1981). Contingency arrangements in personalized instruction. Resources in Education, ED 210 030

Sheehy, M.B. (1990). The effectiveness of Personalized System of Instruction sections versus large lecture recitation classes on the achievement and attitudes of intermediate algebra students. Dissertation Abstracts, 51, 1149A.

Sherman, T.L., Hassett, M.J. & Thompson, R.B. (Eds.) (1977). Innovative teaching methods in introductory college mathematics. Tucson: University of Arizona .

Smith, H.W. (1987). Comparative evaluation of three teaching methods of quantitative techniques: Traditional lecture, Socratic dialogue, and PSI format. Journal of Experimental Education, 55, 149-154.

Smith, J.E. (1976). A comparison of the traditional method and a Personalized System of Instruction in college mathematics. Dissertation Abstracts, 37, 904A.

Smythe, N. (1974). The Keller Plan. The Australian Mathematical Society Gazette, 1, 65-69.

Spencer, R.E. & Semb, G. (1978). Giving students the opportunity to increase unit size: A performance based system for personalized instruction. Journal of Personalized Instruction, 3,76-80

Steele, W.F. (1974) Mathematics 101 at Heidelberg College: PSI vs. tradition. Paper presented at the National Conference on Personalized Instruction, Washington DC.

Stice, J.E. (1975). Seventeen PSI projects at the University of Texas at Austin. Resources in Education, ED 108 665.

Struik, R.R. & Flexner, R.J. (1977). Self-paced calculus: A preliminary evaluation. American Mathematical Monthly, 84, 129-134.

Studer, M. (1976). A personalized system of instruction. The Mathematics Teacher, 69, 381-382, 387.

Svensson, L. (1977). On qualitative differences in learning III: Study skill and learning. British Journal of Educational Psychology, 47, 233-243.

Swetz, F.J., Langguling, H. & Johar, A.R. (1983). Attitudes towards mathematics and school learning in Malaysia and Indonesia: Urban-rural and male-female dichotomies. Comparative Education Review, 27, 394-402.

Taveggia, T.C. (1976). Personalized instruction: A summary of comparative research. American Journal of Physics, 44, 1028-1033.

Taylor, V. (1977). Individualized calculus for the "life long" learner: A two semester comparison of attitudes and effectiveness. Paper presented at the Fourth National Conference of the Center for Personalized Instruction, San Francisco.

Thompson, S.B. (1980). Do individual mastery and traditional instructional systems yield different course effects in college calculus? American Educational Research Journal, 17, 361-375.

Urban, J.D. (1972). The design and evaluation of an undergraduate mathematics course incorporating team teaching, individualized instruction, and team supervision. Dissertation Abstracts, 32, 6301A.

Urbatsch, T.D. (1980). An interaction study: Locus of control, mathematical attitude, mathematical anxiety and mathematical achievement with three modes of instruction: Personalized System of Instruction, traditional lecture, and five day lecture. Dissertation Abstracts, 40, 6180A.

Van Damme, J. & Masui, C. (1980). Locus of control and other student characteristics in interaction with the Personalized System of Instruction. Resources in Education ED 189 145.

Van der Klauw, C.F. & Plomp, T. (1974) Individualized study systems in theory and practice. Higher Education, 3, 213-230.

Van Hiele, P. (1986). Structure and insight: A theory of mathematics education. Orlando FL: Academic Press.

Vatanavigkit, S. (1985). A personalized system of instruction in a mathematics course for Thai college students. Dissertation Abstracts, 46, 370A.

Watkins, D. (1982). Identifying the study process dimensions of Australian university students. Australian Journal of Education, 26, 76-85.

Watkins, D. (1984). Student perception of factors influencing tertiary learning. Higher Education Research and Development, 3, 33-50.

Watkins, D. (1986). The approaches to learning of Australian tertiary students: A replication. Higher Education Research and Development, 5, 185-190.

Watson, J.M. (1978). A Keller Plan variation in Tasmania. The Australian Mathematical Society Gazette, 5, 44-50.

Watson, J.M. (1984). Student choice of assessment in an individualized mathematics course. Programmed Learning and Educational Technology, 21, 34-39.

Watson, J.M. (1986). The Keller Plan, final examination and long-term retention. Journal for Research in Mathematics Education, 17, 60-68.

Wertsch, J.V. & Toma, C. (1995). Discourse and learning in the classroom: A sociocultural approach. In L.P. Steffe & J. Gale Constructivism in education, (pp. 159-174). Hillsdale NJ: Erlbaum

White, M.E. (1974). Differential equations by PSI. In R.S.Ruskin & S.F. Bono (Eds.) Personalized instruction in higher education, (pp. 169-171). Washington DC: Center for Personalized Instruction.

Wolleat, P.L., Pedro, J.D., Becker, A.D. & Fennema, E. (1980). Sex differences in high school students' causal attributions of performance in mathematics. Journal for Research in Mathematics Education, 11, 355-366.

Appendix 1 Control test and achievement scoring

Year 1 examination marks: group comparisons

Results from the controlling analysis of variance involving Year 1 examination marks are documented immediately below in Table 8.1.

Table 8.1. Year 1 examination marks: analysis of variance by teaching group, gender and citizenship.

A. Means

Australians	Women	Men
Keller		
Chem93	62.90 (41)	60.20 (86)
Non-Keller		
Civ93	65.25 (44)	59.21 (193)
Chem92	61.64 (33)	57.18 (94)

Overseas	Women	Men
Keller		
Chem93	66.24 (17)	60.78 (18)
Non-Keller		
Civ93	65.00 (6)	59.83 (35)
Chem92	59.73 (11)	59.47 (17)

Table 8.1 (continued)

B. Analysis of variance

Source of variation	Sum of square	df	Mean square	F	Significance
Main effects					
Teaching group	689.85	2	344.93	1.61	.20
Gender	2133.88	1	2133.88	9.95	.002
Citizenship	65.26	1	65.26	.304	.58
2-way interactions					
Group x gender	149.17	2	74.59	.35	.71
Group x citizenship	31.31	2	15.66	.07	.93
Gender x citizenship	3.55	1	3.55	.02	.90
3-way interactions					
Group x gender x citizenship	142.70	2	71.35	.33	.72
Residual	125050.33	583	214.50		

Examples to illustrate SOLO classification

Question 1

Relational, lower: Angela.

$$y' - \frac{x^2}{x^3 + 1} y^{37} = 0$$

$$\frac{dy}{dx} = \frac{x^2}{x^3 + 1} y^{37}$$

$$\int \frac{1}{y^{37}} dy = \int \frac{x^2}{x^3 + 1} dx$$

$$-\frac{1}{36y^{36}} = \frac{1}{3} \ln (x^3 + 1) + C$$

Multistructural, upper.

1. Bernard

$$y' - \frac{x^2}{x^3+1} y^{37} = 0$$

$$\frac{dy}{dx} - \frac{x^2}{x^3+1} y^{37} = 0$$

$$\frac{dy}{dx} = \frac{x^2}{x^3+1} y^{37}$$

$$\frac{dy}{y^{37}} = \frac{x^2}{x^3+1} dx$$

$$\int y^{-37}\, dy = \int \frac{x^2}{x^3+1}\, dx$$

$$\frac{y^{-36}}{-36} = \int \frac{x^2}{x^3+1}\, dx$$

$$= x^2 \ln|x^3+1| - \int \ln|x^3+1| \,.\, 2x\ dx \quad \text{[wrong]}$$

$$y^{36} = \frac{1}{-36(x^2 \ln|x^3+1| - \int \ln|x^3+1|\ 2x\ dx)}.$$

---

2. Christopher.

$$y' - \frac{x^2}{x^3 + 1} y^{37} = 0$$

$$\frac{dy}{dx} - \frac{x^2}{x^3 + 1} y^{37} = 0$$

$$\frac{1}{y^{37}} dy = \frac{x^2}{x^3 + 1} y^{37} \frac{1}{y^{37}}.$$

$$\frac{1}{y^{37}} dy = \frac{x^2}{x^3 + 1} dx$$

$$\frac{1}{37y^{36}} \ln | y^{37} | + \text{const1} = \frac{1}{3} \ln | x^3 + 1 | + \text{const2}$$ [wrong]

$$\frac{1}{y^{36}} \ln | y^{37} | = 12\ 1/3 \ln | x^3 + 1 | + 37C$$

Multistructural, lower: Daniel

$$\frac{dy}{dx} - \frac{x^2}{x^3 + 1} y^{37} = 0$$

$$\frac{dy}{dx} = \frac{x^2}{x^3 + 1} y^{37}$$

$$\int dy = \int \frac{x^2}{x^3 + 1} y^{37}$$

$$y = y^{37} \int \frac{x^2}{x^3 + 1} d x = \frac{1}{3} y^{37} \int \frac{3x^2}{x^3 + 1} dx$$

[wrong]

$$y = \frac{1}{3} y^{37} \ln | x^3 + 1 | + C$$

Unistructural: Edward

$$y' - \frac{x^2}{x^3 + 1} y37 = 0$$

must arrange this equation in a form separable so can integrate easily

$$\frac{dy}{dx} = \frac{x^2}{x^3 + 1} y^{37} = \frac{x^2/y^3}{x^3/y^3 + 1/y^3} \quad y^{37}$$

$$= \frac{x^2/y^2}{x^3/y^3 + 1/y^3} \quad y^{36} = \frac{(1/v)^2}{(1/v)^3}$$

Question 2

Relational, lower: Francis

(a) u(t)

```
        1|_________________
_______|_________________
          0
```

u(1-t) = u(-t + 1)

If -t < -1 <=> t > 1 => 0

If -t > 1 <=> t < 1 => 1

```
_____1|_____
______|_____|__________________
       0         1
```

r(t) = u(t) u(1-t)

```
      1|_____
       |
_____|____________________
      0         1
```

(b) $y'' - 3y' + 2y = r(t) = u(t)\,u(1 - t)$

$$s^2 L(y) - y'(0) - 3s\,L(y) + 2L(y) = \int_0^{\infty} e^{-st}\, r(t)\, dt$$

$$L(y)\,(s^2 - 3s + 2) = 1 + \int_0^{1} e^{-st}\, 1\, dt + \int_1^{\infty} e^{-st}\, r(t)\, dt$$

$$= 1 + \left[-1/s\ e^{-st} \right]_0^1 + 0$$

$$= 1 - \frac{e^{-s}}{s} + \frac{1}{s}$$

So $L(y) = \dfrac{\{1 - \frac{e^{-s}}{s} + \frac{1}{s}\}}{s^2 - 3s + 2}$

$$= \frac{1 - e^{-s}}{s(s-1)(s-2)} + \frac{1}{(s-1)(s-2)}$$

for this term used the initial condition of y'(0) = 1

(c) using partial fractions

$1 = A(s-1)(s-2) + Bs(s-2) + Cs(s-1)$

$1 = A(s^2 - 3s + 2) + B(s^2 - 2s) + C(s^2 - s)$

$2A = 1 \qquad A = 1/2$

$1/2 + B + C = 0$

$-3/2 - 2B - C = 0$

$-3 - B = 0$ [slip: 3/2 mistaken for 3 1/2]

$B = -3$

$1/2 - 3 + C = 0 \qquad C = 2\ 1/2$

So $\dfrac{1}{s(s-1)(s-2)} = \dfrac{1/2}{s} - \dfrac{3}{s-1} + \dfrac{2\ 1/2}{s-2}$

$L^{-1}\{\frac{1/2}{s} - \frac{3}{s-1} + \frac{2\ 1/2}{s-2}\} = L^{-1}(\frac{1}{2s}) - L^{-1}(\frac{3}{s-1}) + L^{-1}(\frac{2\ 1/2}{s-2})$

$= 1/2 - 3\,e^{t} + 2\ 1/2\,e^{2t}$

also need $L^{-1}\{-e^{-s}\left(\frac{1}{2s} - \frac{3}{s-1} + \frac{2\,1/2}{s-2}\right)\}$

$$= -\{L^{-1}\left(\frac{-e^{-s}}{2s}\right) - L^{-1}\left(\frac{3e^{-s}}{s-1}\right) + L^{-1}\left(\frac{2\,1/2\ e^{-s}}{s-2}\right)\}$$

$$= -(1/2\,u(t-1) - 3e^{t-1}u(t-1) + 2\,1/2\ e^{2(t-1)}u(t-1)\}$$

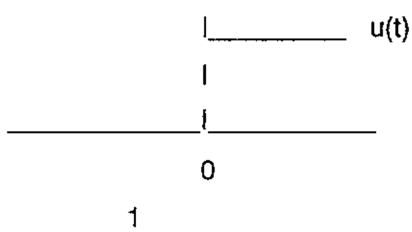
So all together now:

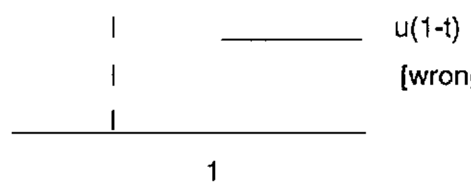
$$L^{-1}\left\{\frac{1-e^{-s}}{s(s-1)(s-2)}\right\} = 1/2 - 3e^{t} + 2\,1/2\,e^{2t} - 1/2\,u(t-1) + 3e^{(t-1)}u(t-1)$$
$$- 2\,1/2\,e^{2(t-1)}u(t-1)$$

where u(t) is the Heaviside step function.

Multistructural, upper: Gina.

(a) $r(t) = u(t)\,u(1-t)$





[wrong]

r(t) [consistent]

(b) $y'' - 3y' + 2y = u(t)\,u(1-t)$ $y(0) = 0$ $y'(0) = 1$

L{ u(t) u(1-t) } = L{ u(t) u(-(t-1)) } = (1 - e-s)/s [no reasons given]

taking Laplace transforms both sides

$$s^2 Y(s) - s\,y(0) - y'(0) - 3s\,Y(s) - y(0) + 2\,Y(s) = L\{ u(t)\,u(1-t) \}$$

$$s^2 Y(s) - 1 - 3(s\,Y(s)) + 2\,Y(s) = L\{ u(t)\,u(1-t) \}$$

$$(s^2 - 3s + 2)\,Y(s) = 1 + L\{ u(t)\,u(t-1) \}$$

$$(s-1)(s-2)\,Y(s) = 1 + \frac{1 - e^{-s}}{s}$$

$$Y(s) = \frac{1}{(s-1)(s-2)} + \frac{1 - e^{-s}}{s(s-1)(s-2)}$$

Term corresponding to the initial conditions is $\dfrac{1}{(s-1)(s-2)}$.

(c) $L^{-1}\left\{ \dfrac{1 - e^{-s}}{s(s-1)(s-2)} \right\}$

$$\text{Let } F(s) = \frac{1 - e^{-s}}{s(s-1)(s-2)} = \frac{1}{s(s-1)(s-2)} - \frac{e^{-s}}{s(s-1)(s-2)}$$

$$= H(s) - G(s)$$

$$= H(s) - e^{-s} H(s)$$

by partial fractions,

$$\frac{1}{s(s-1)(s-2)} = \frac{A}{s} + \frac{B}{s-1} + \frac{C}{s-2}.$$

$$1 = A(s-1)(s-2) + Bs(s-2) + Cs(s-1)$$

s = 0: $1 = A(-1)(-2) = 2A \qquad A = 1/2$

$s = 1$: $\quad 1 = B(-1) \quad B = -1$

$s = 2 \quad 1 = C(2)(1) \quad C = 1/2$

$$\frac{1}{s(s-1)(s-2)} = \frac{1}{2s} - \frac{1}{s-1} + \frac{1}{2(s-2)}. = H(s)$$

$$L^{-1}\{H(s)\} = h(t) = 1/2 - e^{t} + 1/2\, e^{2t} \qquad (A)$$

$$G(s) = \frac{e^{-s}}{s(s-1)(s-2)} = e^{-s} H(s)$$

where $\quad L^{-1}\{e^{-s} H(s)\} = h(t-1)\, u(t-1)$

from (A) $\quad h(t-1) = 1/2 - e^{t-1} + 1/2\, e^{2(t-1)}$

therefore

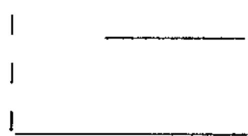
$$L^{-1}\{e^{-s} H(s)\} = \{1/2 - e^{t-1} + 1/2\, e^{2(t-1)}\}\, u(t-1)$$

so

$$L^{-1}\left(\frac{1 - e^{-s}}{s(s-1)(s-2)}\right) = 1/2 - e^{t} + 1/2\, e^{2t} - \{1/2 - e^{t-1} + 1/2\, e^{2(t-1)}\}\, u(t-1)$$

Multistructural, lower: Henry

(a)



[wrong]

(b) $L\{r(t)\} = s^2 Y - 1 - 3s Y + 2Y = L\{u(t)\, u(1-t)\}$

$$Y(s^2 - 3s + 2) = 1 + L\{u(t)\, u(1-t)\}$$

$$\frac{1}{s^2 - 3s + 2} = Y + Y\,L\{ u(t)\, u(1-t) \}$$ [wrong]

$$Y = \frac{1}{s^2 - 3s + 2} + L\{ u(t)\, u(1-t) \}$$ [wrong]

$$= \frac{1}{(s-1)(s-2)} + L\{ u(t)\, u(t-1) \}$$ [wrong]

$$= \frac{1}{(s-1)(s-2)} + \frac{1 - e^{-s}}{s(s-1)(s-2)}$$ [wrong]

(c) $$\frac{1 - e^{-s}}{s(s-1)(s-2)} = \frac{a}{s} + \frac{b}{s-1} + \frac{c}{s-2}$$ [wrong]

$$1 - e^{-s} = a(s-1)(s-2) + bs(s-2) + cs(s-1)$$

$$1 - e^{-s} = as^2 - 3as + 2a + bs^2 - 2bs + cs^2 - cs$$

equate like coefficients

$2a = 1$ $\quad a = 1/2$

$-3a - 2b - c = -1$ [wrong]

$a + b + c = 0$

$-2b - c = -2\ 1/2$ [wrong]

$b + c = -1/2$

$-2b + 1/2 + b = -2\ 1/2$

$-2b + b = -3$

$b = 3$ $\quad c = -3\ 1/2$

$$\frac{1 - e^{-s}}{s(s-1)(s-2)} = \frac{1}{2s} + \frac{3}{s-1} - \frac{7}{2(s-2)}$$

$$L^{-1}\left\{\frac{1 - e^{-s}}{s(s-1)(s-2)}\right\} = 1/2 + 3e^{t} - 7/2\, e^{2t}$$

Unistructural: Iris

$y'' - 3y' + 2y = r(t)$ $y(0) = 0$ $y'(0) = 1$

$r(t) = u(t)\,u(1-t)$

```
           I_______________ u(t)                              I          u(1-t)
           I                                                  I
   _______I________________                      ___________I______________
                                                              I
                                                 ___________I -1
```

```
        I
        I               r(t)
        I_______________
```

Appendix 2 Attitudes: instrument and details of results

Instrument: Explanation, background information, questionnaires

Mathematics student survey

The questionnaires that follow are part of a study of how mathematics students approach their work in mathematics, and how they feel about the organisation of their mathematics subjects.

The answer sheets should be completed in pencil, coding in numbers as usual, and marking responses to questions by filling in the appropriate space as follows ( ).

The mathematics subject we are interested in for this group is the main mathematics subject you did last year. [Initial version: appropriately changed for the second data collection]. For most people this is Mathematics 1 (MATH1032/1042), but for some it will be a second year subject, possibly MATH2021 or MATH2009. Please answer with the appropriate subject in mind.

Please start by putting identification and background details on the answer sheet. We need students' names so that later we can correlate answers with subject marks. The material will be used purely for research purposes, and it is most unlikely that individual answers will be looked at even by the research team, because the answer sheets will be machine read, so please speak freely in your questionnaire answers.

For these, please fill in the spaces on the answer sheets that correspond most closely to your attitude or your usual behaviour.

Background details

Name_______________________________________

Surname Given names

Student number___________________

Sex M ( ) F ( )

Age________

Citizenship: Australian ( )

Other ( )
(Specify)_______________

Main mathematics subject last year__________________
[this year in follow up]

Questionnaires

Approaches to study questionnaire

1. Working on my mathematics for this subject is just a boring task.

2. I find it hard to get organised in this mathematics subject.

3. I expect to get a good grade in this mathematics subject.

4. How I do in mathematics is up to me.

5. I get quite worried about how I will do in this mathematics subject.

6. Once I start working on my mathematics problems, I find it hard to stop.

7. I try to test my understanding of a mathematics topic by doing lots of examples.

8. If I find a mathematics topic hard, I tend to panic.

9. I usually feel in control of my work in this mathematics subject.

10. I can't seem to get interested in this mathematics subject.

11. I am sure I will pass this mathematics subject.

12. It's most unusual for me to be far behind in this mathematics subject.

13. I have learned to take responsibility for my own work in mathematics.

14. I enjoy the challenge of a new topic in mathematics.

15. I feel very confident about handling the work in this mathematics subject.

16. I tend to put off work in this mathematics subject.

17. I don't expect to do very well in this mathematics subject.

18. It's important to me to do as well as I can in this mathematics subject.

19. I find this mathematics subject quite easy.

20. I feel overwhelmed by the work in this mathematics subject.

21. When I study mathematics I have trouble settling down to work.

22. When I do a mathematics test, I think I do worse than I should, because I tend to get anxious about it.

23. I really try to understand the work we do in this mathematics subject.

24. I do some work on this mathematics subject every week.

<u>Organisation of teaching questionnaire</u>

This questionnaire is about the organisation of mathematics subjects. Please tick the space on the answer sheet that most closely corresponds to your attitude to the subject we are asking about.

25. In this mathematics subject, we get enough feedback on how we are going.

26. I would like more help than I am getting in this mathematics subject.

27. If I did another mathematics subject, I would be happy for it to be organised like this one.

28. In this mathematics subject, the way it is organised doesn't help me to work for real understanding.

29. I have learned a lot in this mathematics subject.

30. In this mathematics subject the workload is too heavy.

31. The way this mathematics subject is organised makes enough allowance for individual differences.

32. Among the mathematics subjects I have done, this is one of the most effectively taught.

Details of analyses involving attitudes.

Comparisons involving gender

The factorial analyses of covariance by teaching group (with non-Keller groups combined), gender and citizenship group, with Year 1 examination scores as covariate are documented immediately below, in Tables 9.1 to 9.6. Results indicate that the main analyses could safely omit the gender factor, thus helping to overcome the difficulty of very small cell sizes when all factors were used.

Table 9.1. Year 2 confidence scores: comparisons by teaching group (with combined non-Keller groups), gender and citizenship, controlled for Year 1 examination scores.

A. Means

Group(s)	Women	Men
Keller		
Australians	3.39 (37)	3.30 (72)
Overseas	3.40 (14)	3.37 (17)
Non-Keller		
Australians	3.39 (45)	3.42 (159)
Overseas	3,52 (7)	3.40 (37)

B. Analysis of covariance

Source of variation	Sum of squares	df	Mean Square	F	Significance
Covariate					
Y1 exam	19.12	1	19.12	48.53	.000
Main effects					
Teaching group	.68	1	.68	1.72	.56
Gender	.04	1	.04	.10	.75
Citizenship	.04	1	.04	.10	.76
2-way interactions					
Group x gender	.15	1	.15	.38	.54
Group x citizenship	.01	1	.01	.03	.89
Gender x citizenship	.01	1	.01	.01	.91
3-way interaction					
Group x gender x citizenship	.13	1	.13	.32	.58
Residual	149.36	379	.44		

Table 9.2. Year 2 study methods scores: comparisons by teaching group (with combined non-Keller groups), gender and citizenship, controlled for Year 1 examination scores.

A. Means

Group(s)	Women	Men
Keller		
Australians	3.71 (36)	3.43 (69)
Overseas	3.60 (13)	3.63 (17)
Non-Keller		
Australians	3.17 (45)	3.21 (160)
Overseas	3,52 (7)	3.40 (37)

B. Analysis of covariance

Source of variation	Sum of squares	df	Mean Square	F	Significance
Covariate					
Y1 exam	1.03	1	1.03	3.06	.08
Main effects					
Teaching group	6.59	1	6.59	9.78	.000
Gender	.17	1	.17	.50	.48
Citizenship	2.11	1	2.11	6.29	.01
2-way interactions					
Group x gender	1.31	1	1.31	3.89	.05
Group x citizenship	.20	1	.20	.61	.44
Gender x citizenship	.42	1	.42	1.26	.26
3-way interaction					
Group x gender x citizenship	.14	1	.14	.41	.52
Residual	126.04	375		.34	

Table 9.3 Year 2 motivation scores: comparisons by teaching group (with combined non-Keller groups), gender and citizenship, controlled for Year 1 examination scores.

A. Means

Group(s)	Women	Men
Keller		
Australians	3.38 (37)	3.17 (71)
Overseas	3.38 (14)	3.44 (17)
Non-Keller		
Australians	3.01 (45)	2.87 (160)
Overseas	3.14 (7)	3.32 (37)

B. Analysis of covariance

Source of variation	Sum of squares	df	Mean Square	F	Significance
Covariate					
Y1 exam	2.68	1	2.68	5.80	.02
Main effects					
Teaching group	6.30	1	6.30	13.65	.000
Gender	.57	1	.57	1.23	.27
Citizenship	5.59	1	5.59	12.12	.001
2-way interactions					
Group x gender	.15	1	.15	.31	.58
Group x citizenship	.37	1	.37	.80	.37
Gender x citizenship	.99	1	.99	2.14	.14
3-way interaction					
Group x gender x citizenship	.00	1	.00	.01	.94
Residual	174.94	379	.46		

Table 9.4. Year 2 anxiety scores: comparisons by teaching group (with combined non-Keller groups), gender and citizenship, controlled for Year 1 examination scores.

A. Means

Group(s)	Women	Men
Keller		
Australians	2.94 (34)	2.96 (70)
Overseas	2.95 (14)	2.71 (17)
Non-Keller		
Australians	3.34 (43)	3.23 (159)
Overseas	3.05 (7)	3.08 (37)

B. Analysis of covariance

Source of variation	Sum of squares	df	Mean Square	F	Significance
Covariate					
Y1 exam	11.97	1	11.97	30.69	.000
Main effects					
Teaching group	8.61	1	8.61	22.06	.000
Gender	.03	1	.03	.08	.48
Citizenship	1.65	1	1.65	4.24	.04
2-way interactions					
Group x gender	.04	1	.04	.09	96
Group x citizenship	.01	1	.01	.03	.44
Gender x citizenship	.06	1	.06	.14	.71
3-way interaction					
Group x gender x citizenship	.45	1	.45	1.32	.28
Residual	145.13	372	.39		

Table 9.5. Year 2 control scores: comparisons by teaching group (with combined non-Keller groups), gender and citizenship, controlled for Year 1 examination scores.

A. Means

Group(s)	Women	Men
Keller		
Australians	4.03 (37)	3.86 (71)
Overseas	4.07 (14)	4.12 (17)
Non-Keller		
Australians	3.56 (45)	3.70 (161)
Overseas	3.86 (7)	3.73 (37)

B. Analysis of covariance

Source of variation	Sum of squares	df	Mean Square	F	Significance
Covariate					
Y1 exam	2.87	1	2.87	4.77	.03
Main effects					
Teaching group	5.99	1	5.99	9.95	.01
Gender	04	1	.04	.07	.79
Citizenship	.85	1	.85	1.42	.23
2-way interactions					
Group x gender	.93	1	.93	1.54	.22
Group x citizenship	10	1	.10	.17	.68
Gender x citizenship	.00	1	.00	.00	.99
3-way interaction	.66	1	.66	1.09	.52
Group x gender x citizenship					
Residual	228.65	380	.60		

Table 9.6. Year 2 subject perception scores: comparisons by teaching group (with combined non-Keller groups), gender and citizenship, controlled for Year 1 examination scores.

A. Means

Group(s)	Women	Men
Keller		
Australians	3.18 (32)	3.26 (65)
Overseas	3.32 (13)	3.26 (16)
Non-Keller		
Australians	3.19 (44)	3.28 (156)
Overseas	3.18 (7)	3.04 (35)

B. Analysis of covariance

Source of variation	Sum of squares	df	Mean Square	F	Significance
Covariate					
Y1 exam	1.85	1	1.85	5.86	.02
Main effects					
Teaching group	.06	1	.06	.20	.66
Gender	.24	1	.24	77	.38
Citizenship	.58	1	.58	1.03	.18
2-way interactions					
Group x gender	.00	1	.00	.00	.99
Group x citizenship	.52	1	.52	1.66	.20
Gender x citizenship	.32	1	.52	1.02	.31
3-way interaction					
Group x gender x citizenship	.02	1	.02	.06	.81
Residual	113.15	359	.32		

Comparisons by teaching group and citizenship

For completeness, the full analyses of covariance, factored by the three teaching groups and the two citizenship groups, are given here, in Tables 9.7 to 9.12. The results are outlined in Chapter 4.

Table 9.7. Year 2 confidence scores: comparisons by teaching group and citizenship, controlled for Year 1 examination scores.

A. Means

Group	Australians	Overseas
Keller		
Chem93	3.33 (109)	3.39 (31)
Non-Keller		
Civ93	3.37 (144)	3.36 (28)
Chem92	3.52 (61)	3.52 (16)

B. Analysis of covariance

Source of variation	Sum of squares	df	Mean Square	F	Significance
Covariate					
Y1 exam	19.12	1	19.12	49.50	.000
Main effects					
Teaching group	2.95	2	1.47	3.81	.02
Citizenship	.02	1	.02	.04	.84
2-way interactions					
Group x citizenship	.03	2	.02	.04	.96
Residual	147.56	382	.39		

Table 9.8. Year 2 study methods scores: comparisons by teaching group and citizenship, controlled for Year 1 examination scores.

A. Means

Group	Australians	Overseas
Keller		
Chem93	3.52 (105)	3.62 (30)
Non-Keller		
Civ93	3.28 (145)	3.51 (28)
Chem92	3.01 (61)	3.34 (16)

B. Analysis of covariance

Source of variation	Sum of squares	df	Mean Square	F	Significance
Covariate					
Y1 exam	1.03	1	1.03	3.11	.08
Main effects					
Teaching group	10.10	2	5.05	15.26	.000
Citizenship	2.36	1	2.36	7 .12	.01
2-way interactions					
Group x citizenship	.51	2	.26	.77	.46
Residual	125.05	378	.33		

Table 9.9. Year 2 motivation scores: comparisons by teaching group and citizenship, controlled for Year 1 examination scores.

A. Means

Group	Australians	Overseas
Keller		
Chem93	3.24 (108)	3.41 (31)
Non-Keller		
Civ93	2.89 (145)	3.31 (28)
Chem92	2.91 (61)	3.27 (16)

B. Analysis of covariance

Source of variation	Sum of squares	df	Mean Square	F	Significance
Covariate					
Y1 exam	2.69	1	2.69	5.77	.02
Main effects					
Teaching group	7.42	2	3.71	7.96	.000
Citizenship	5.70	1	5.70	12.23	.001
2-way interactions					
Group x citizenship	.81	2	.41	.87	.42
Residual	177.93	382	.47		

Table 9.10. Year 2 anxiety scores: comparisons by teaching group and citizenship, controlled for Year 1 examination scores.

A. Means

Group	Australians	Overseas
Keller		
Chem93	2.96 (104)	2.82 (31)
Non-Keller		
Civ93	3.19 (144)	3.01 (28)
Chem92	3.42 (61)	3.18 (16)

B. Analysis of covariance

Source of variation	Sum of squares	df	Mean Square	F	Significance
Covariate					
Y1 exam	11.97	1	11.97	31.51	.000
Main effects					
Teaching group	12.16	2	6.08	16.01	.000
Citizenship	1.87	1	1.87	4.92	.03
2-way interactions					
Group x citizenship	.10	2	.05	.14	.87
Residual	142.39	375	.38		

Table 9.11. Year 2 control scores: comparisons by teaching group and citizenship, controlled for Year 1 examination scores.

A. Means

Group	Australians	Overseas
Keller		
Chem93	3.92 (108)	4.10 (31)
Non-Keller		
Civ93	3.74 (146)	3.82 (28)
Chem92	3.51(61)	3.63 (16)

B. Analysis of covariance

Source of variation	Sum of squares	df	Mean Square	F	Significance
Covariate					
Y1 exam	2.87	1	2.87	4.82	.03
Main effects					
Teaching group	8.27	2	4.14	6.95	.001
Citizenship	.97	1	.97	1.62	.20
2-way interactions					
Group x citizenship	10	2	.05	.08	.92
Residual	228.11	383	.60		

Table 9.12. Year 2 subject perception scores: comparisons by teaching group and citizenship, controlled for Year 1 examination scores.

A. Means

Group	Australians	Overseas
Keller		
Chem93	3.24 (97)	3.28 (29)
Non-Keller		
Civ93	3.30 (141)	3.07 (26)
Chem92	3.17 (60)	3.05 (16)

B. Analysis of covariance

Source of variation	Sum of squares	df	Mean Square	F	Significance
Covariate					
Y1 exam	1.85	1	1.85	5.90	.02
Main effects					
Teaching group	.49	2	24	.77	.46
Citizenship	.52	1	.52	1.67	.20
2-way interactions					
Group x citizenship	.79	2	.39	1.25	.29
Residual	177.93	382	.47		

Attitude changes over Year 2

Existence of changes in attitude scores

Attitude scores from the beginning and end of Year 2 were compared using paired t-tests, to check whether attitude changes were significantly different from zero. Results for teaching groups and citizenship groups are given below in Tables 9.13 and 9.14

Table 9.13 Changes in attitudes over Year 2: separate teaching groups

A. Chem93

Variable	Mean (difference)	Standard Deviation	t	df	Significance
Confidence	.09	.69	1.53	128	13
Study methods	.20	.57	3.84	125	.000
Motivation	.36	.66	6.15	129	.000
Anxiety	-.23	.64	-3.99	125	.000
Control	.34	.90	4.32	128	.000
Subject	.42	.78	4.24	61	.000

B. Civ93

Variable	Mean (difference)	Standard Deviation	t	df	Significance
Confidence	.30	.66	5.42	145	.000
Study methods	-.06	.56	-1.33	144	.19
Motivation	16	.57	3.46	145	.001
Anxiety	.13	.64	2.47	144	.02
Control	.06	.89	.83	146	.41
Subject	.33	.63	6.04	136	.000

Table 9.14 Changes in attitudes over Year 2: separate citizenship groups

A. Australian

Variable	Mean (difference)	Standard Deviation	t	df	Significance
Confidence	.22	.71	4.68	216	.000
Study methods	.02	.59	.47	213	.64
Motivation	.27	.62	6.45	216	.000
Anxiety	-.03	.66	-.60	212	.55
Control	.19	.94	2.97	215	.003
Subject	.37	.67	7.09	163	.000

B. Overseas

Variable	Mean (difference)	Standard Deviation	t	df	Significance
Confidence	.11	.60	1.42	57	.16
Study methods	.21	.52	3.03	55	.004
Motivation	.19	.63	2.30	57	.03

Anxiety	-.08	.65	-.98	56	.33
Control	.19	.74	1.96	57	.06
Subject	26	.72	2.11	34	.04

Attitude changes: group comparisons

Attitude change scores over Year 2 were tested for group differences using factorial analysis of covariance by teaching group and citizenship, controlled for Year 1 examination scores. Results are summarised in Chapter 4 , and full details are given below in Tables 9.15 to 9.20.

Table 9.15. Year 2 changes in confidence scores: comparisons by teaching group and citizenship with control for Year 1 examination marks.

A. Means

Group	Australians	Overseas
Keller: Chem93	.09 (96)	.08 (31)
Non-Keller: Civ93	.33 (117)	.19 (25)

B. Analysis of covariance

Source of variation	Sum of squares	df	Mean Square	F	Significance
Covariate					
Y1 exam	.24	1	.24	.52	.47
Main effects					
Teaching group	2.96	1	2.96	6.34	.01
Citizenship	.20	1	.20	.43	.51
2-way interactions					
Group x citizenship	.18	1	.18	.39	
Residual	123.95	264	.47		

Table 9.16. Year 2 changes in study methods scores: comparisons by teaching group and citizenship with control for Year 1 examination marks

A. Means

Group	Australians	Overseas
Keller: Chem93	16 (93)	.33 (30)
Non-Keller: Civ93	-.09 (117)	.06 (25)

B. Analysis of covariance

Source of variation	Sum of squares	df	Mean Square	F	Significance
Covariate					
Y1 exam	.01	1	.01	.03	.89
Main effects					
Teaching group	4.44	1	4.44	13.74	.000
Citizenship	1.28	1	1.28	3.98	.05
2-way interactions					
Group x citizenship	.00			.01	.91
Residual	83.95	260	32		

Table 9.17 Year 2 changes in motivation scores: comparisons by teaching group and citizenship with control for Year 1 examination marks

A. Means

Group	Australians	Overseas
Keller: Chem93	.38 (96)	.23 (31)
Non-Keller: Civ93	.18 (117)	.06 (25)

B. Analysis of covariance

Source of variation	Sum of squares	df	Mean Square	F	Significance
Covariate					
Y1 exam	.73	1	.73	1.94	.17
Main effects					
Teaching group	3.00	1	3.00	8.00	.01
Citizenship	.36	1	.36	.96	.33
2-way interactions					
Group x citizenship	.01	1	.01	.02	.90
Residual	99.12	264	38		

Table 9.18. Year 2 changes in anxiety scores: comparisons by teaching group and citizenship with control for Year 1 examination marks

A. Means

Group	Australians	Overseas
Keller: Chem93	-.24 (92)	-.23 (31)
Non-Keller: Civ93	13 (117	.10 (25)

B. Analysis of covariance

Source of variation	Sum of squares	df	Mean Square	F	Significance
Covariate					
Y1 exam	.43	1	.43	1.07	.30
Main effects					
Teaching group	8.15	1	8.15	20.14	.000
Citizenship	.00	1	.00	.00	.96
2-way interactions					
Group x citizenship	.02	1	.02	.05	.83
Residual	105.23	260	.41		

Table 9.19 Year 2 changes in control scores: comparisons by teaching group and citizenship with control for Year 1 examination marks

A. Means

Group	Australians	Overseas
Keller: Chem93	.31 (95)	.45 (31)
Non-Keller: Civ93	.12 (118)	-.08 (25)

B. Analysis of covariance

Source of variation	Sum of squares	df	Mean Square	F	Significance
Covariate					
Y1 exam	.32	1	.32	.40	.53
Main effects					
Teaching group	4.74	1	4.74	6.01	.02
Citizenship	.00	1	.00	.00	.96
2-way interactions					
Group x citizenship	.29	1	.29	1.65	.20
Residual	20741	264	.79		

Table 9.20 Year 2 changes in subject perception scores: comparisons by teaching group and citizenship with control for Year 1 examination marks

A. Means

Group	Australians	Overseas
Keller: Chem93		.39 (13)
Non-Keller: Civ93		.18 (22)

B. Analysis of covariance

Source of variation	Sum of squares	df	Mean Square	F	Significance
Covariate					
Y1 exam	2.46	1	2.46	5.38	.02
Main effects					
Teaching group	.73	1	.73	1.60	.21
Citizenship	.15	1	.15	.32	.57
2-way interactions					
Group x citizenship	.20	1	.20	.44	.51
Residual	87.23	191	.46		

Appendix 3 Interview schedule and details of results

Interview schedule

The interview questions are reprinted immediately below

Mathematics student project: interviews

Name

Sex Age Citizenship

Degree enrolled for

1. Where did you grow up?

2. Did you go to secondary school in Australia?

3. Do you have an Australian Higher School Certificate?

4. Did you do Mathematics 1 at the University of New South Wales?

5. We want mainly to find out how your second year mathematics went. First, though, could you tell me about your general attitude to your whole university mathematics work.

a) Would you have chosen to do some mathematics at university?

b) Are you interested in mathematics?

c) Do you find it easy to study mathematics actively and regularly?

6. Could you tell me a bit about your mathematics studies in [name of Year 2 mathematics subject]

a) Did you manage to keep up to date?

b) Did you find it interesting?

c) Did you feel confident about the work?

d) Did you feel very responsible for keeping up with your work in mathematics?

e) Did you feel that you really got to understand what you studied?

7. a) Did your study methods in [subject] change compared with how they were in your previous mathematics subjects?

b) In particular, did you do more work or less?

c) Did you work more steadily during the year than you used to?

8. a) Did you feel the workload in your mathematics was heavy, reasonable or light?

b) What effect did your workload over all subjects have on your mathematics study?

9. Did you feel the degree of independence required of you was too high, about right or too low?

10. How do you think your understanding of the mathematics in Year 2 would compare with that of your previous work in mathematics?

a) Better or worse than in previous mathematics subjects?

b) Do you remember it any better?

11. Keller group: In your Year 2 mathematics you worked independently from notes. Did you feel that, compared with lectures, this was better, worse or about the same?

Non-Keller group: How do you feel about the traditional lecture-tutorial organisation of mathematics teaching?

12. What did you feel about the method of assessment in your Year 2 mathematics subject?

a) Was the level of performance required good for you?

b) Did you feel the assessment was fair?

c) What effect did it have on your work methods?

d) Did it make you more or less interested in the subject?

e) Was your confidence lower or higher?

13. a) Did you feel you got enough help from the teaching staff?

b) Enough support in general? Who from?

14. a) Did you feel overall that the organisation of the subject was a help, a hindrance, or no worse than usual?

b) Would you say it was better suited to some students than others?

Interview results

The interview results are outlined below. Sample background is included. It is to be noted that the sample was very small, and largely Australian, but included overseas students in two teaching groups. That is, the sample contained 9 chemical engineering students from the Keller Plan subject, and 7 other engineering students from the traditionally taught subject. Interview material was collected after students had had experience of the teaching method for at least three quarters of an academic year.

Results are summarised immediately below.

Table 10.1 Interview results: Background data

A. Gender

	Keller	Non-Keller
Women	4	4
Men	5	3

B. Age

	Keller	Non-Keller
18		1
19	6	3
20	2	3
22	1	

C. Citizenship

	Keller	Non-Keller
Australia	6	6
Malaysia	2	1
Sri Lanka	1	

Table 10.2. Interview results: Questions

1. Country of upbringing

	Keller	Non-Keller
Australia	4	6
Malaysia	2	1
New Zealand	1	
Sri Lanka	2	

2. Secondary schooling

	Keller	Non-Keller
Australia	4	6
Malaysia	2	1
New Zealand	1	
Sri Lanka	2	

3. University entrance

All had Australian qualifications.

4. Year 1 mathematics

	Keller	Non-Keller
UNSW	9	5
Other Australian university		2

5. Mathematics in general

a) Would have chosen

	Keller	Non-Keller
Y	5	3
N	3	3

Previously Y, now N		1

b) Interested

	Keller	Non-Keller
Y	8	1
A little bit		4
N	1	2

c) Studies easily

	Keller	Non-Keller
Y	5	3
Fairly	3	4
N	1	2

6. Year 2 mathematics

a) Up to date

	Keller	Non-Keller
Y	7	2
N	2	5

b) Interested

	Keller	Non-Keller
Y	4	1
Fairly	3	4
N	2	2

c) Confident

	Keller	Non-Keller
Y	4	4
Fairly	1	3
Not very	2	
N	2	

d) Responsible

	Keller	Non-Keller
Y	7	5
Tutor helped		1
N		1

e) Understanding

	Keller	Non-Keller
Good	6	2
Fair	3	3
Poor		1

7. Study methods

a) Changed this year

	Keller	Non-Keller
Y	7	2
N	3	4

b) More work

	Keller	Non-Keller
Y	7	2
Better	1	
Same		4
Less	1	1

c) Steadier work

	Keller	Non-Keller
Y	3	2
Same	2	1
More flexible	1	
Less	2	4

8. Workload

a) Mathematics

	Keller	Non-Keller
About right	5	3
A bit heavy	3	3
Heavy		1

b) Total

	Keller	Non-Keller
A bit heavy	1	1
Very heavy	8	5

Three Keller Plan students said their mathematics suffered from their general overload.

9. Level of independence

	Keller	Non-Keller
Good	6	4
Satisfactory	2	
Too high	1	1
Too low		1

10. Understanding

a) Compared with other mathematics subjects

	Keller	Non-Keller
Better	5	2
Harder	1	
Same	2	3
Worse		1
Cram and forget	1	

b) Remember

	Keller	Non-Keller
Better	5	3
Same	2	3
Worse	2	1

11. Different questions for the two groups

Keller : How does Keller Plan compare with lectures?

	Keller
Better	5
Okay	1
Okay if more solved problems	1
Hard	1
Worse	1

Non-Keller: How do you feel about the normal lecture-problem class-tutorial organisation of your mathematics subject?

	Non-Keller
Printed outlines wanted	1
More examples wanted	1
More tutorials, no lecture	1
Tutorials okay; lectures a waste	1
Tutorials and problem classes okay, not lectures	1
Don't know	1

12. Assessment method

a) Level required

	Keller	Keller
Good for one	4	
Good but hard	2	
Satisfactory		6
Too high	3	

b) Fair

	Keller	Non-Keller
Y	3	5
Moderately good		2
Uncertain	1	
Markers vary; some too strict	5	

c) Did it affect study?

Keller

	Keller
More careful	2
More comprehensive	3
More determined	1
More independent	1
Not much affected	1

Non-Keller

	Non-Keller
Regular assignments helped	4
Crammed	1
Not much affected	2

d) Did it affect interest?

	Keller	Non-Keller
More interest	2	
A bit more	2	2
Same	2	2
Less	3	3

e) Did it affect confidence?

	Keller	Non-Keller
Higher	5	
Okay		3
Same	1	2
One exam makes you anxious		2
Lower	3	

13. Enough help?

a) From staff

	Keller	Non-Keller
Y	5	2
Sometimes		1
Didn't seek help	2	3
N	2	1

b) From others

	Keller	Non-Keller
Friends	7	5
Didn't seek help	2	1
Not enough		1

14. Subject organisation

a) Did it help?

	Keller	Non-Keller
y	6	
Okay		2
Okay if good teachers		1
Yes but administrative problems	1	
Average		4
Made it a struggle	1	
Disadvantage	1	

b) Was it better for some?

	Keller	Non-Keller
Better for well organised motivated students.	7	4
Same for most	1	2
Don't know		1
Not good for any	1	