Psychology of Learning: improving pupil performance

4

MASTERING LEARNING

Although early attempts at developing a scientific approach to learning were not always as successful as their proponents anticipated, they have led to an approach to learning that does produce consistent benefits in terms of achievement, benefits that are educationally significant.

Mastery learning is embodied in Bloom's (1968) 'Learning For Mastery' and Keller's (1968) 'Personalized System of Instruction.' Bloom's approach was developed to exploit mastery learning in the schoolroom, whereas Keller developed his system for higher education. However, both have been applied in many different contexts and have been found to be very powerful methods for increasing student performance in a wide range of activities.

BLOOM'S LEARNING FOR MASTERY

Bloom (1968) indicated that a learning strategy for mastery may be derived from the work of Carroll (1963), with support from the ideas of Morrison (1926), Skinner (1954) and Bruner (1966). Block (1971) traces the influences back to Washburne's (1922) Winnetka Plan and Morrison's approach at the University of Chicago's Laboratory School of 1926. In both systems mastery was defined in terms of particular educational objectives and mastery of each unit was required of students before they proceeded to the next unit. An ungraded diagnostic-progress test was administered at the completion of each unit to provide feedback on the adequacy of the student's learning. The test either indicated mastery, usually set at a level of 80-90%, or it highlighted the material the student still had to master. On the basis of the results of the diagnostic test each student's original instruction could be supplemented with appropriate remedial materials so that he could complete his unit, by obtaining a score reaching the mastery criterion. Although popular well into the 1930's these ideas eventually disappeared due primarily to a lack of the technology to sustain them (Block, 1971). They didn't emerge again until the late 1950's and 1960's, as a corollary of programmed instruction. The basic idea underlying programmed instruction being that the learning of any behaviour, no matter how complex, rested upon the learning of a sequence of less complex component behaviours. Theoretically, by breaking down complex behaviour into a chain of component parts and by ensuring mastery of every link in the chain, it would be possible for any student to master even the most complex skills.

The research in programmed instruction showed that it worked very well for some students, especially those requiring small steps, drill and frequent reinforcement, but it was not so effective for all students, and it did not provide a useful mastery model in itself.

Bloom (1968) based his 'Learning For Mastery' (LFM) on the model proposed by J. Carroll (1963). This requires 'that the task can be unequivocally described and that means can be found for making valid judgement as to when the learner has accomplished the learning task - that is, has achieved the learning goal which has been set for him.'

A Model of School Learning

Carroll made the assumption that the work of the school can be broken down into a series of learning tasks. Although he admitted that this can be called into question because, in the school, the various tasks to be learned are not necessarily treated as being separate and distinct, and the process of teaching is often organised so that learning will take place 'incidentally' and in the course of other activities.

Some activities are not included as being appropriate for the model of learning that Carroll has in mind. For example, it is not intended to apply to those goals of the school which do not lend themselves to being considered as learning tasks, such as those having to do with attitudes and dispositions.

The model says that the learner will succeed in learning a given task to the extent that he spends the amount of time that he needs to learn the task. 'Spending time' means actually spending time on the act of learning. It is not 'elapsed time', but the time during which the person is oriented to the learning task and actively engaged in learning.

The variables in the model are considered under two headings: 1) determinants of time needed for learning, and 2) determinants of time spent in learning.

Figure 4-1. Aptitude-Achievement for Uniform Instruction per Learner.

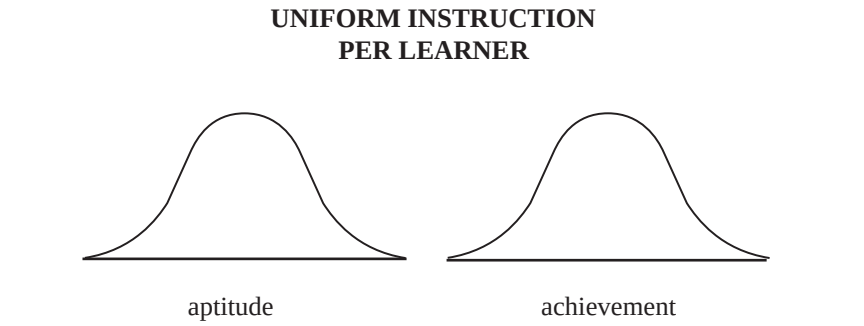
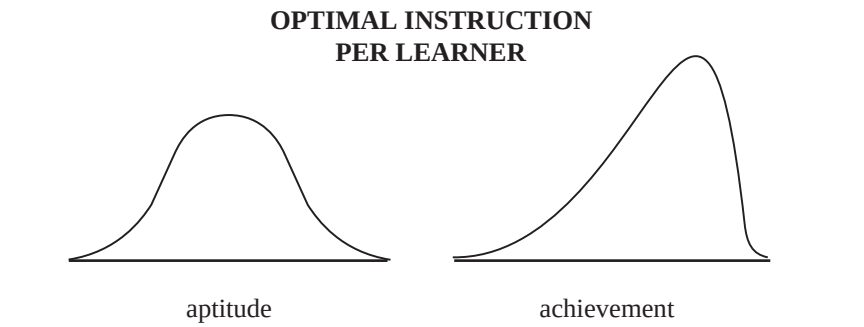


Figure 4-2. Aptitude-Achievement for Optimal Instruction per Learner.



Time Needed in Learning

There are three factors affecting the time needed in learning, according to Carroll: Aptitude, Quality of Instruction, and Ability to Understand Instruction.

1. Aptitude

Common experience, as well as abundant research evidence, suggests that the amounts of time needed by the children even under ideal conditions will differ widely. Carroll suggests that the amount of time each pupil will need to learn a task under optimal conditions is the primary measure of a variable he calls 'aptitude for learning this task.' Learners who need only a small amount of time are said to have high aptitude; learners who need a large amount of time are said to have low aptitude. Some learners, Carroll admits, will never learn even under these optimal conditions; and these learners would need an indefinitely large (or an infinite) amount of time to learn the task.

The measure of aptitude is specific to the task under consideration. It may be regarded as a function of numerous other variables, such as the amount of prior learning or it may also depend upon a series of traits or characteristics of the learner which enter into a wide variety of tasks. These traits may be accounted for solely on the basis of generalized prior learnings, or they may reflect genetically determined individual characteristics.

Bloom (1968) contrasts this view with a common view that high levels of achievement are possible only for the most able students. Implicit in Carroll's formulation is the idea that all students can achieve mastery of a given task if they are provided with sufficient time. Therefore, mastery is available to all, if the right means can be found to help each individual. Bloom quotes evidence to support this, based on the grade norms for many standardized achievement tests. Such norms show that certain scores attained by the most able students at one grade level are achieved by the majority of students at a later grade.

Bloom believes that there is likely to be a small percentage of students (1%-5%), with special talents for a given subject, who are able to learn and use a subject with greater fluency than other students. These may be inherited, or may be developed by special training or interests. There is another small percentage of individuals who have extreme difficulty in learning particular subjects. Between these extremes are approximately 90% of the individuals for whom, Carroll

and Bloom believe, aptitudes are predictive of rate of learning rather than the level or complexity of learning that is possible.

This means that 95% of students are capable of learning a subject to a high level of mastery, with up to 95% of a class being capable of attaining the equivalent of an A-grade, previously associated with the top 5-10%. Bloom assumes that for some students more time, effort and help may be required to reach this level, and for some the time and help required may make such achievement prohibitive.

Thus, a major problem for mastery-learning strategists is to find a way of reducing the time and effort that the less able students are required to make, to a point where it is not prohibitive. Bloom points out that students' aptitudes can be changed by appropriate learning experiences in the school and in the home, and that such changes can be most effectively brought about in the elementary years of schooling and during the early years in the home. Even if such changes cannot be effected in the early years it is highly probable that more effective learning conditions can reduce the amount of time required for learning to mastery.

2. Quality of Instruction

Carroll states that one job of the teacher is to organize and present the task to be learned in such a way that the learner can learn it as rapidly and as efficiently as he is able. This means that the learner must be put into adequate sensory contact with the material to be learned and that the various aspects of the learning task must be presented in such an order and with such detail that every step of the learning is adequately prepared for by a previous step. It may also mean that the instruction must be adapted for the special needs and characteristics of the learner, including his stage of learning. This variable applies not only to the performance of a teacher but also to the characteristics of textbooks, workbooks, films, and teaching programmes.

If the quality of instruction is anything less than optimal, it is probable that the learner will need more time to learn the task than he would otherwise need, although some learners will be more handicapped by poor instruction than others.

Bloom (1968) is concerned that we have fallen into the trap of looking at the class of 30 or so students and asking the questions 'What is the best teacher for the group?'; 'What is the best method of instruction for the group?'; 'What is the best instructional material for the group?' He maintains that it is necessary to start from a very different perspective in which each individual student may require very different types and qualities of instruction. In other words 'the same content and objectives of instruction may be learned by different students as the result of very different types of instruction.' For example, some students may need more examples, some may require more approval or reinforcement, or more concrete illustrations or explanations.

The example of home tutoring of middle-class students is quoted by Bloom to illustrate the benefits of the quality of instruction suited to the student's needs. A third of the middle-class students in one study received home tuition in algebra when parents saw the school as failing to provide the necessary tuition. These students received relatively high grades for algebra, with the relationship between mathematics aptitude tests at the start of the course and their achievement at the end of the year being almost zero. With no extra tuition the relationship between aptitude and achievement was highly correlated (+0.9).

3. Ability to Understand Instruction

A further factor affecting the time needed in learning is the ability to understand instruction. Carroll suggests that this variable interacts with the method of instruction in a special and interesting way. It could be measured as some combination of 'general intelligence' and 'verbal ability'; the former of these two would come into play 'in instructional situations where the learner is left to infer for himself the concepts and relationships inherent in the material to be learned', while the latter would come into play 'whenever the instruction utilized language beyond the grasp of the learner.'

Bloom comments that in schools there is frequently a single teacher and a single set of instructional materials for all students in a given class. The student who has facility in understanding the teacher's communications has little difficulty in learning the subject. However, the student with the same basic aptitude for the subject but lacking the ability to understand the teacher or the text will experience considerable difficulty in learning.

This is the point at which the student's abilities interact with the instructional materials. For Bloom this ability will primarily be determined by verbal ability and reading comprehension because schools place such great stress on these abilities.

Bloom (1964) in 'Stability and Change in Human Characteristics' presents evidence that it is possible to produce changes in verbal ability early in life by appropriate training, but that such changes are more difficult later on in life, although vocabulary and reading ability may be improved to some extent at all ages. Therefore, the greatest immediate pay-off in this area will come from modifications in the instruction. Bloom suggests that given help and various types of aids, individual teachers can find ways of modifying their instruction to fit the differing needs of their students. He gives examples of how to overcome this problem within the school context:

1) **Group Study:** Small groups of students meeting regularly to go over points of difficulty were found to be most effective, especially if this could be arranged in a non-competitive atmosphere, with more able students being provided with opportunities to

strengthen their understanding while helping less able students to grasp ideas presented through alternative ways of explanation.

2) **Tutorial Help:** This is very costly, but should be made available to students as they require it, especially where difficulties cannot be corrected in any other way.
3) **Text Books:** The provision of various text books providing instruction for different abilities, rather than adopting one set book, can obviously be of benefit.
4) **Workbooks and Programmed Texts:** Drill and practice or small steps with frequent reinforcement may help some students and programmed texts or workbooks can provide these.
5) **Audio-visual Methods and Academic Games:** Bloom suggests that some students may learn a particular idea from concrete illustrations and vivid explanations which AV methods (video tapes, filmstrips etc.) can provide. Other students may require direct laboratory experiences, games and computer simulations. In other words the highest priority need not always be given to abstract and verbal ways of instruction. The suggestion is not for particular materials to be used by particular students throughout the course, but for such materials as are appropriate to be used in helping students to overcome their difficulties.

Time Spent in Learning

In Carroll's model there are two factors which directly affect the time which the student actually spends in learning.

1. Time Allowed for Learning or 'Opportunity'

This is the first factor in the time actually spent in learning. Carroll comments that schools may allow less than adequate time for many students learning many tasks. Schools respond to differences in learning rates in many ways, in particular, they ignore these differences; a certain amount of time is provided for everybody to learn, and no more. At the other extreme is the case where each student is allowed to proceed exactly at his own rate; private instruction in music or foreign languages and self-instruction by teaching-machine or other means are approximations to this case. The middle position is occupied by learning situations in which there is some kind of 'ability grouping.'

Obviously, failure to allow enough time for learning produces incomplete learning. If a person needs two hours to learn something and is allowed only one hour, and if we assume that learning proceeds linearly with time, the degree of learning is only 50%. Carroll feels that one of the most aversive things which a school can do is not to allow sufficient time for a well-motivated child to master a given learning task before the next step is taken up. Children meet such frustrations by indifference or the more extreme avoidance reactions and are, in any case, handicapped in undertaking the next task.

This is the key to mastery. Aptitude and other factors determine the rate of learning and all students can achieve mastery if they devote the amount of time needed, and for this they must be provided with sufficient time.

However, it is not just the amount of time that accounts for the level of achievement, but the amount of time in relation to other factors such as quality of instruction, aptitude and verbal ability and also the student's ability to persevere with the task in hand.

2. Perseverance

The second factor is the amount of time the learner is willing to spend in the learning process. Carroll describes this as consisting of three attitudes: (1) a marked willingness to spend time beyond the ordinary schedule, in a given task; (2) a willingness to withstand discomfort, which includes adjusting to shortened lunch hours, or no lunch hours, working without holidays and withstanding fatigue; (3) a willingness to face failure, which brings with it a realization that patient work may lead to successful termination of the task in hand.

A learner who needs a certain amount of time to learn a task may or may not be willing to persevere for that amount of time in trying to learn. The learner may not be motivated to learn at all, or may regard the task as something too difficult for him to learn; in either case, he may spend no time at all in trying to learn. He may, of course, be so highly motivated that he would be willing to spend more time than he needs in order to reach a specified criterion of mastery.

This variable, if it is not sufficiently great to allow the learner to attain mastery, operates in the conceptual model to reduce the degree of learning. It is a function partly of 'motivation' or the desire to learn and may also be a function of what are ordinarily called emotional variables. For example, one may desire to learn but be unable to endure frustrations caused by difficulties in the learning task.

Students approach different tasks with different degrees of perseverance and, if frustrated in learning, will reduce the time devoted to a subject. Bloom suggests that it may be difficult to manipulate the perseverance of a student and that changes in the type of instruction and learning materials may help students to master material for a given level of perseverance. Then, as a corollary of this action, the reward and perceived success can raise

the student's level of perseverance. As students attain mastery they are likely to raise their level of perseverance in related areas.

Bloom found that with appropriate instructional resources the demands for perseverance may be reduced. Frequent feedback accompanied by specific help in instruction and materials can reduce not only the time taken but also the perseverance required. Improvement in the quality of instruction (explanations, illustrations, etc.) may also reduce the required perseverance. He concludes that:

> There seems to be little reason to make learning so difficult that only a small proportion of the students can persevere to mastery. Endurance and unusual perseverance may be appropriate for long-distance running - they are not great virtues in their own right. The emphasis should be on learning, not vague ideas of discipline and endurance. (Bloom, 1968/81, p.164)

Thus, the model involves five elements - three residing in the individual and two stemming from external conditions. Factors in the individual are (1) aptitude - the amount of time needed to learn the task under optimal conditions, (2) ability to understand instruction, and (3) perseverance - the amount of time the learner is willing to engage actively in learning. Factors in external conditions are (4) opportunity - the time allowed for learning, and (5) the quality of instruction - a measure of the degree to which instruction is presented so that it will not require additional time for mastery beyond that required in view of aptitude.

Assessment Criteria

Carroll's model advocates a 'criterion referenced' assessment procedure, as suggested by Tyler (1932), in which the student's performance is judged according to how well he has done by comparison with some predetermined criterion. This contrasts with the traditional 'norm referenced' assessment in which a given student is judged against the performance of his colleagues.

Criterion referenced tests provide a standard against which all student performances can be measured, and can ensure a minimum standard of competence. The driving test is a good example of this kind of assessment: all learners take the test and if they don't reach the criterion level they fail but may repeat the test until they reach the set level of competence. Norm referencing of a group at a driving school may result in a good distribution of grades from fail to A-grade, but this would not necessarily mean that even the A-grade students were able to reach the criterion level. It would mean that they could out-perform their colleagues, but the overall standard could still be below that required for safe driving on the highway.

The programmed learning movement popularized criterion referencing, with the programme writer producing a criterion test related to the objectives of his programme. The aim of the programme was to teach in such a way that 90% of the students obtained 90% on the criterion test. To attain this the programme writer would test the programme and revise areas of the text which produced low scores in the test. He would have to:

> resist the pressure from psychometricians who have had him insert some extra-hard questions and cut out easy ones - 'in order to spread students out a bit and show that some have achieved more than others.' He would reply, 'No, my intention is not to show whether some have learned more than others but to sow whether or not they have learned well enough to have attained the objectives. If all score 100% I shall be well pleased.' (Rowntree, 1977, p.180)

With norm referenced tests the aim is to produce the sort of achievement distribution shown in Figure 4-1, whereas the criterion referenced test should produce the achievement results shown in Figure 4-2.

Bloom (1968) took these ideas and transformed Carroll's conceptual model into an effective working model for mastery. He argued that if aptitudes were predictive of the **rate** at which (and not the **level to** which) a student could learn a given task, then it should be possible to fix the degree of learning expected of students at some level of mastery and to systematically manipulate the variables in Carroll's model so that all or almost all students attained it. He also reasoned that if aptitude for a subject was normally distributed in the student population and they were provided with uniform instruction in terms of quality and learning time, then achievement on completion of instruction would be normally distributed (Figure 4-1).

However, if the learner were to receive optimal quality of instruction and learning time then a majority of students could expect to attain mastery and there would be little or no relationship between aptitude and achievement (Figure 4-2).

Learning Hierarchies

Bloom also incorporated Gagné's theory of learning hierarchies in his learning for mastery approach. This theory states that the learning of any intellectual skill is important because it supports the learning of more complex skills. The support is only there if the previously learned skills are readily available when the new learning is taking place. Gagné suggests that mastery means 'readily accessible to recall at the time of learning of the more complex skill.'

Learning hierarchy research started with Gagné's early work with programmed learning of mathematical skills. In the first investigation a programmed book was used to teach a hierarchy of twenty-two elements leading to the ability to solve linear equations. Gagné felt that each element in the hierarchy could not be acquired unless the learners possessed those prerequisite skills found at lower levels in the hierarchy. These prerequi-

sites could be determined by asking 'What would the individual have to be able to do in order that he can attain successful performance on this task, provided he is given only instructions?' (Gagné, 1962, p.358). He proposed that transfer from one learning set to another standing above it in the hierarchy will be zero if the lower one cannot be recalled, and will be as much as 100% if it can be recalled. From this it follows that attaining each new learning set is dependent on recall of previous subordinate sets interacting with the effects of instructions.

> In order for learning to occur at any point in the hierarchy, according to this theory, each of the learning sets subordinate to a given task must be highly recallable, and integrated by a thinking process into the solution of the problem posed by the task. The attainment of the final task is thus conceived to be a matter of successive attainment and "integration" of a series of lower level learning sets, beginning with those which are already available to the learner. (Gagné and Paradise, 1961, p.2)

The results tended to support Gagné's theory and the numbers of subjects who learned the higher elements without possessing the relevant subordinate ones were small but not zero, as they should have been if the hierarchy was correct (Gagné and Paradise, 1961). A similar result, but with fewer exceptions to the postulated hierarchy, was found in the next experimental investigation of mathematical subject matter (Gagné, Mayor, Garstens and Paradise, 1962).

Despite the importance of the implications of Gagné's early work, few investigations of the validity of hierarchies were performed in the subsequent decade (White, 1973). Those that did take place did not always confirm the existence of the postulated hierarchy. For example, Kolb (1967-8) found that many subjects passed tests of higher elements dealing with basic ideas of ratio and line-segment graphs, having failed at the lower ones. Although in later experiments subjects behaved in ways which did not fit the proposed hierarchies and consequently brought the theory into disrepute, White (1973) suggested that the studies were at fault because of their unsatisfactory design. Bergan (1980), however, disagreed with White and suggested that the results arose from the limitations in the learning hierarchy itself. One limitation is that Gagné's model does not include a set of equations describing the relationships among variables in the model. Such equations could be beneficial in estimating the magnitude of influence of variables in the model. The major limitation seems to be the restriction of the description of the relationships depicted in the model to two conditions: a skill is either prerequisite or it is not, no other possibilities are allowed. This fails to take account of the varieties of relationships that could exist.

> There could be reciprocal causal relationships between skills on the same level in a hierarchy. Unspecified variables outside the model could be exerting a significant influence on skill learning. There could be relationships among variables that have no simple causal interpretation. Finally, a skill placed in a hierarchy by virtue of one prerequisite relationship could influence several other skills at different levels in the hierarchy without being prerequisite to any of them. (Bergan, 1980, p.628)

There is nevertheless great appeal in the idea of learning hierarchies and it is an integral part of Bloom's approach to learning for mastery. Bloom (1976) does, however, acknowledge that not all learning tasks contain hierarchical arrangements of elements.

Learning For Mastery Strategies

Bloom (1968) acknowledged the fact that there were many different strategies for mastery learning, but that each strategy should deal effectively with the variables described by Carroll (1963).

He suggested that one ideal strategy would be to have a good tutor for each individual student, but acknowledged that this is not possible on a large scale on economic grounds. Nevertheless, the tutor-student relationship is a useful model to consider when working out the details of a more economically viable strategy.

Other strategies suggested include allowing students to go at their own pace, guiding students as to the most suitable courses, and providing different tracks or streams.

Bloom and his colleagues, following in Morrison's footsteps at the University of Chicago, experimented with various mastery strategies using diagnostic procedures and alternative methods and materials as supplements to regular classroom instruction. With such an approach the aim was to bring all students up to mastery standard within the regular term or period of calendar time in which the course was usually taught.

The strategy required specification of the 'pre-conditions' necessary, the development of 'operating procedures' and evaluation of the 'outcomes' of the strategy.

Pre-conditions

To develop mastery learning it is necessary, according to Bloom, to define what mastery is and whether or not a student has achieved it. Specification of objectives and content of instruction are, therefore, necessary preconditions. The translation of these into evaluation procedures is also necessary. Implicit in defining outcomes and evaluation procedures is the distinction between the teaching-learning process and the evaluation process. The teaching-learning is intended to prepare a student for achievement of a specific level of competence in an area of learning. The summative evaluation at the end of a course is intended to determine the extent to which this has been achieved. This is not a competitive approach to

education because Bloom sees much of learning and development being destroyed by primary emphasis on competition.

The LFM approach relies on setting standards of mastery and excellence, followed by a strategy to bring as many students as possible to this standard. One method employed was to base the standards for the new approach on those achieved by the top group prior to the introduction of the mastery approach. In this way students were informed of the performance required but were not in competition for grades, they were to be judged on the basis of levels of mastery actually obtained by students in the previous year. This enabled a more co-operative approach, students helping each other without the fear that special advantage was being given.

Operating Procedures

Bloom's mastery approach was based on materials and methods used in previous years, within a similar timetable, in the belief that requiring extensive training of teachers would limit the acceptance of the new approach.

The operating procedures were used as supplements to the regular teaching. They were devised to ensure mastery of each learning unit in such a way as to reduce time required by altering quality of instruction and ability to understand the instruction. Two main operating procedures were adopted:

1. Formative and Summative Evaluation

The course was broken down into small units of learning, such as a chapter in a textbook, a time unit of the course or a well defined portion of a course, usually a week or two of learning activity.

Based on the ideas of Bloom (1956) and Gagné (1965) each unit was analysed into a number of elements ranging from specific terms or facts, to more complex abstract ideas such as concepts or principles and finally to applications of principles and the analysis of complex theoretical statements. Much of this initial work was conducted with such subjects as algebra and science topics, which fitted in well with Gagné's theoretical position on hierarchies of learning tasks.

Following from this breaking down of subject matter brief diagnostic progress tests were devised, which were then used to determine the extent of student mastery of a given unit. Such tests also indicated what the student still had to accomplish when he failed to reach the mastery criterion. These diagnostic tests were referred to as Formative Evaluation.

Frequent formative evaluation paces the students and helps motivate them, according to mastery theory, ensuring that each unit of learning is mastered before subsequent learning tasks are started. Early parts of the course may require more frequent formative testing than later sections, but there will be some variation in testing in various parts of a course.

Bloom believed that, for students who thoroughly mastered the unit, formative evaluation would reinforce the learning and reduce course anxiety about achievement.

Students lacking mastery of a particular unit had the specific deficits identified and particular instructional materials or processes were recommended to correct the learning. A very specific prescription was seen as being essential for students failing a formative test.

The test was not graded other than as showing mastery or non-mastery because Bloom believed that giving different point-grades (A, B, C, etc.) could be counter-productive.

Thus, the main purpose of the formative evaluation was to determine the degree of mastery of a given learning task and to identify those aspects of the task which were not mastered.

> The purpose is not to grade or certify the learner; it is to help both the learner and the teacher focus upon the particular learning necessary for movement toward mastery. (Bloom, Hastings and Madaus, 1971, p.61)

The term 'summative evaluation' was given to the much more general assessment of the degree to which the 'larger outcomes have been attained over the entire course.' Such summative evaluation could be used to grade pupils and to report to parents or administrators. It was felt that several skills or concepts which combine to make a broader ability should be presented before a summative examination is administered. Summative tests should not be thought of only as end of course or final examinations, they could equally be used every four or six weeks to assess mid-term attainment of the various abilities taught in a course.

Bloom, Hastings and Madaus (1971) suggest that it is the level of generalization which differentiates summative from formative evaluation, and they indicate that summative assessment is linked to Tyler's more general descriptions of desired behaviours, whereas formative evaluation is more closely allied to Gagné's detailed prerequisite capabilities for each large aim.

2. Alternative Learning Resources

Bloom found that students do attempt to work on their difficulties if they are given specific suggestions of what to do. The best procedure identified was to have small groups of students (two or three) meet regularly for up to an hour each week to review the results of formative tests and to help each other overcome difficulties. Tutorial assistance was found not to be as popular as peer group meetings for secondary and higher education students. Other types of learning resource prescriptions, recommended by Bloom, were:

1. Rereading pages of the original instructional materials.
2. Reading alternative materials.
3. Use of specific pages of workbooks or programmed texts.

4. Use of selected a-v aids.

Ryan and Schmidt (1979) identified the most successful corrective strategies as being those which include objectives plus a problem testing the objectives of the previous lesson, discussion of the problem, specific prescriptions for using the text, class notes and handouts, and alternative resources, such as texts, workbooks, games and kits. When correctives consisted of objectives or problems only their effectiveness was considerably diminished.

Affective and Cognitive Consequences of LFM

Bloom suggested that if a system of formative and summative evaluation informs the student of his mastery of the subject, he will come to believe in his own mastery and competence and this will change his view of himself and the world.

He will certainly begin to respond more positively to the subject he has mastered. Mastery will provide the necessary reassurance and reinforcement to help the student to view himself as adequate thus contributing to positive mental health:

> Finally, modern society requires continual learning throughout life. If the schools do not promote adequate learning and reassurance of progress, the student must come to reject learning - both in the school and later in life. Mastery learning can give zest to school learning and can develop a life-long interest in learning. It is this continual learning which should be the major goal of the educational system. (Bloom, 1968/81, p.174)

Bloom's initial description of the cognitive outcomes described results for an experiment conducted from 1965 to 1967. Using the standards of grading adopted in 1965, prior to the introduction of a mastery approach, the aim was to use a variety of techniques during ensuing years to turn all students into 'A' students (ie. all students in the years 1966 and 1967 would obtain the same grades, using the same criteria, as were obtained by the top students of 1965). The results indicated that for 1967 80% of students obtained the level of mastery previously associated with 'A' grade students and by 1967 this had risen to 90% of students. The techniques certainly appeared to work in these initial studies and led to further investigations.

Learning For Mastery: The Research

Bloom (1976) summarized the early research on the mastery approach in 'Human Characteristics and School Learning.' These were 'a series of small experimental and school-based studies' in which students from the University of Chicago contrasted mastery with traditional methods.

Cognitive entry behaviours were measured at the beginning of each learning task, using a formative diagnostic test, and students were tested again at the end of the teaching session. Those students in the mastery groups received additional tuition, as indicated by the formative diagnostic test, if they failed to reach the set mastery criterion. They were then tested again after these 'corrective procedures' using a parallel form of formative test.

If Bloom's theory is correct we would expect a student's performance on one learning task to be highly correlated with performance on the next task, in the traditional group. The best students in this group would perform well on the first task and this would enable them to perform at an equally high level on the next task.

This would not be the case in the mastery group because those students who were failing to reach mastery on the initial task would receive appropriate materials and time to enable them to reach a position previously associated with the more able students. Having gained this position they would be just as likely to perform at a higher level in the next task. Thus, their performance on the original task would not be predictive of their performance on the second task. That at least is the theory, and it also predicts that there would be high correlations between one task score and the next task score for traditional courses, but much smaller (or even zero) correlations between one task and the next for mastery groups.

Bloom (1976) quotes studies covering subjects such as second language acquisition, matrix algebra, probability and an 'imaginary' science. The correlations between achievement on the first learning task in the series and the final (summative) achievement are:

	Mastery	**Non-mastery**
Anderson (1973)	0.31	0.68
Arlin (1973)	0.19	0.49
Block (1970)	0.44	0.78
Levin (1975)	0.59	0.72
Ozcelik (1974)	0.36	0.48
Median	0.36	0.68

These results demonstrate an overall agreement with Bloom's prediction from his theory, although the mastery correlations do not reach the zero figure that Bloom had hoped for. There is a definite trend for a much reduced correlation between initial performance and final performance under mastery conditions. The larger correlation of 0.68 for non-mastery students indicates the extent to which performance on the initial learning task will influence the final achievement when there are no correctives for defective learning.

Of further interest is the average cumulative performance of the five experimental groups reported by Bloom (1976) given in Table 4.1. The non-mastery group does not receive correctives and, therefore, is represented by one score only for each of the three hierarchically arranged learning tasks. This is called the 'Original Achievement' score. The mastery group receives correctives after each learning task, for those

Table 4-1. Mean performance scores for Mastery and Non- mastery groups.

	LEARNING TASK 1		LEARNING TASK 2		LEARNING TASK 3		
	Orig. Ach.	Corr. Ach.	Orig. Ach.	Corr. Ach.	Orig. Ach.	Corr. Ach.	Summ. Ach.
Mastery Groups	63	85	69	87	71	86	76
Non-mastery Groups	63		55		54		44

Orig. Ach. - Original achievement after instruction on learning task.
Corr. Ach. - Corrected achievement (mastery group only) after receiving additional corrective instruction.
Summ. Ach. - Summative achievement on completion of three tasks.

Table 4-2. Student achievement in mastery and non-mastery conditions. (Based on Kim, 1971)

	Percentage of Students Achieving Mastery			
	English	**Maths**	**Physics**	**Biology**
Instructional Treatment				
Mastery	50	48	30	22
Non- Mastery	44	26	8	3

failing to reach the criterion level, and takes a parallel form of the test giving a 'Corrected Achievement' score in addition to the 'Original Achievement' score. The difference between the 'Original Achievement' score and the 'Corrected Achievement' score represents the improvement in the mastery group following the correctives procedure. This improvement in performance for the mastery group results in a larger difference between the groups on the next learning task. With each set of correctives we see the gap between the mastery and non-mastery group increasing: initially, there is no difference; on the second task in the hierarchy there is a difference of 14%, in favour of the mastery group; and on the third task this has further increased to 17%. The final summative average difference is 32%, in favour of the mastery group. These results were indicative of the scale of improvements possible with a mastery learning approach, and were confirmed by the research studies reviewed by Block (1974).

Block quotes the work of Kim (1971) and his colleagues (Lee et al., 1971) as showing the power of LFM. In these studies, using a group-based, mastery learning strategy, with pupil-teacher ratios of 70:1, Kim and his colleagues converted much of the Korean elementary and middle-school system into a mastery model. The pilot study taught 272 seventh grade students an eight-session unit on geometry. Kim found that 75% of the mastery group scored at the mastery criteria of 80% correct, compared with only 40% of the traditionally taught students reaching this level.

On the basis of this successful experiment Kim expanded his next study to 5,800 students, who studied maths andEnglish over an eight week period. In English, 75% of mastery students reached the 80% criterion, compared with 28% of the traditional group; in maths, 67% of mastery students reached the criterion, against 39% of the traditional group.

The third experiment was much larger in scope. Rural and urban pupils learned maths, English, physics and biology over a full school year. There were 25,887 students in the experiment. The results were quite impressive (Table 4-2), with 2,000-3,000 more students

gaining mastery in maths, physics and biology when compared with traditional methods.

Similar results were obtained by Lee et al. (1971), for arithmetic and science when taught to 12,504 pupils in 5th and 6th grades, with 27-32% more pupils reaching the 80% criterion under mastery conditions.

These differences in performance demonstrate the scale of achievement possible for students who are not reaching their maximum potential under normal instructional circumstances. Block (1974) applauds the magnitude of the successful application of mastery learning:

> Here a difference of a few percentage points, let alone differences of 20 to 30 points, between the percentage of students reaching a high performance level under mastery conditions and the percentage reaching the same level under non-mastery learning conditions translates into differences of thousands of students who are achieving levels in their learning they might never ordinarily have reached. (Block, 1974, p.34)

Further research on mastery learning in the classroom was summarized by Block and Burns (1977) in 'Review of Research in Education' (vol. 4), edited by L.S. Shulman. In this later paper the authors reported a meta-analysis on the research, giving an average effect size for the LFM-taught students of 0.83, which would move the average student in the non-mastery group from the 50th position in class to the top 20% of the class, a large educational effect. Keller studies were also included in this review and before discussing the overall results a description of Keller's 'Personalized System Of Instruction' is necessary.

KELLER'S PERSONALIZED SYSTEM OF INSTRUCTION

Keller describes the the first steps towards his system as having taken place during an evening brain-storming session involving Keller and three other psychologist colleagues, who had been invited to set up a new department of Psychology in the University of Brazilia, in 1963. They were given absolute freedom to follow their own dictates and fancies in carrying out the task.

Whereas Bloom had adapted the work of Carroll and Gagné, Keller's group was overtly 'operant' in its orientation, or as Keller put it 'tarred with the brush of reinforcement theory.' They were excited by Skinner's Natural Science Course at Harvard University, and all were convinced that traditional teaching methods were hopelessly out-of-date, and were impressed by the teaching machine and programmed-instruction movement.

They were not the only ones disillusioned with traditional methods of instruction. In the late 1950s and early 1960s educational researchers raised some unsettling questions about the effectiveness of most college teaching. In a number of major studies, investigators found that they could reduce by more than two-thirds the amount of time students spent in the classroom without affecting end-of-course student achievement (Kulik and Jaksa, 1977). Dubin & Taveggia (1968) went even further and demonstrated that such typical classroom activities as lectures and discussions might be eliminated completely without altering student achievement. Pooling data from 91 studies conducted between 1924 and 1965 they showed that students learned as much from independent study as they did from more conventional approaches to teaching. In fact, one of the chapters in Dubin and Taveggia's 'The Teaching-Learning Paradox' is titled "THERE ARE NO DIFFERENCES", in very large capital letters.

On the basis of these research results many innovators believed that students could learn as much through the newer approaches as they had learned from the more traditional approaches. And some even believed that they might actually improve learning.

According to the behaviouristic approach of Keller and his colleagues students are more likely to perform well if they find satisfaction in their studies or in behaviouristic terms 'The consequences of learning are important instructional contingencies.'

This means that positive consequences (instructor praise, good grades, feelings of achievement) are considered to be much more effective facilitators of learning than negative consequences (boredom, failure, or other forms of punishment).

External rewards are important, but it was thought that for adult learners they were not nearly as important as the general internal rewards (feelings of achievement, satisfaction or accomplishment).

Keller's 'Personalized System of Instruction' (PSI), which is also known as the 'Keller Plan', is based on these principles. It specifies objectives and provides reinforcement for their successful achievement, as well as giving the student more options and opportunities for personal interaction than traditional instructional systems.

The system was first publicly announced in 1968 in Keller's article 'Goodbye, teacher', which appeared in the new Journal of Applied Behavior Analysis. It begins with the chant:

> Goodbye scholars, goodbye school
> Goodbye teacher, darned old fool

and goes on to describe Keller's reasons for being disillusioned with the traditional classroom and formal education. He also outlines some of his observations made during his time in a military training centre. This is what he says about that centre:

> I should have seen many things that I didn't see at all, or saw very dimly. I could have noted, for example, that instruction in such a centre was highly individualised,

> in spite of large classes, sometimes permitting students to advance at their own speed throughout a course of study. I could have seen the clear specification of terminal skills for each course, together with the carefully graded steps leading to this end. I could have seen the demand for perfection at every level of training and for every student; the employment of classroom instructors who were little more than the successful graduates of earlier classes; the minimizing of the lecture as a teaching device and the maximizing of student participation. I could have seen, especially, an interesting division of labour in the educational process, wherein the non-commissioned, classroom teacher was restricted to duties of guiding, clarifying, demonstrating, testing, grading, and the like, while the commissioned teacher, the training officer, dealt with matters of course logistics, the interpretation of training manuals, the construction of lesson plans and guides, the evaluation of student progress, the selection of non-commissioned cadre, and the writing of reports for his superiors. (Keller, 1968, p.80)

Much of the Personalized System of Instruction can be seen as a development of Keller's observations of military training - perhaps one reason for the emphasis on instruction, rather than 'education' or 'teaching'. There are five main elements in PSI as described in the 1968 paper:

1. The unit-perfection requirement for advance, which lets the student go ahead to new material only after demonstrating mastery of the preceding unit.
2. The go-at-your-own pace feature, which permits a student to move through the course at a speed commensurate with his ability and other demands upon his time.
3. The stress upon the written word in teacher-student communications.
4. The use of proctors, which permits repeated testing, immediate scoring, almost unavoidable tutoring, and a marked enhancement of the personal-social aspect of the educational process.
5. The use of lectures and demonstrations as vehicles of motivation, rather than sources of critical information.

Mastery

Mastery learning proposes that all or almost all students can master what they are taught. It enables 90-95% of students to achieve the same high level as the top 25% learning under typical group-based instructional methods. Under such circumstances it is assumed that students will learn more material in less time and there will be greater student interest in and attitude toward the subject learned than is found in traditional classrooms.

Keller indicates that the mastery requirement means an essentially perfect performance. It is acceptable to interpret this as each student answering nine out of ten questions correctly, provided the single error is corrected. Any further compromise with mastery is unwise because:

> the requirement of eighty-five percent will automatically raise questions as to why eighty-three percent isn't 'close enough.' There will be requests for 'partial credit.' An adversary system with 'bargaining' about grades quickly returns to the classroom. The mutual respect and high morale upon which the course depends deteriorates. The course policy statement should specify that the criterion for a pass is perfection.
> (Keller and Sherman, 1974, p.37)

Self-pacing

One of the essential characteristics of mastery learning is the time spent in study and this leads to the emphasis, in the Keller Plan, placed on a completely student-paced programme of study.

It is important to stress that the system places no restraints on the student's study time - thus a student may elect to take a unit quiz at her convenience or when she is most ready to demonstrate mastery, rather than at a time for all students which is dictated by the instructor.

The self-pacing allows for individual differences in rate of study and is essential if the level of achievement is held constant. This is because in essence it is not possible to hold both level of individual achievement and rate of completion constant. Traditional lecture methods, for example, hold rate of completion constant and, because of this, level of achievement varies widely. However, in PSI courses tests can be taken at any time and can be repeated as required, and the result is that while the average student evidences a steady progress, at any one time individual students will be working on different units and taking different tests. Here we have mastery level of achievement held constant and so the rate of completion varies.

Written Materials

Keller in 'Goodbye, Teacher' summarizes his use of written materials by stating that PSI places stress upon

the written word in teacher-student communication. Although programmed texts are used, Keller emphasizes that the steps in a PSI programme are not frames in a set; they are more inclusive and can be better described as reading assignments or laboratory exercises. Also, he states in a later description of PSI that television, computers and teaching machines may all be used in the system but should not be equated with the system.

PSI courses are based upon a standard textbook, journal articles, other readings and are supplemented by a study guide. This guide breaks the course into units. Each guide consists of an introduction, statement of objectives, procedure, study questions, and supplementary materials as dictated by the specific course.

The study guide is designed to allow the student to work through the material independently and this function over-rides the form the guide might take. A typical guide has four sections:

1. Introduction: in which the instructor comments upon the textbook and supplements it. It may point out differing views to those expressed in the text and correct deficiencies.
2. Statement of objectives: which indicates the goals of the study unit in behavioural terms. It tells the student what he will be able to do on completion of the unit. The objectives should be very specific, but make the work interesting at the same time.
3. Study questions: these should help the student to detect his progress and success, and may include concepts or technical terms that require defining and problems for completion.
4. Procedure: this section specifies the activities necessary for completion of the unit objectives. It may consist only of the assignment to be read, or it may be more detailed, providing an indication of specific pages to be read and the order for working through different articles. A page-and-line key of special instructions may be included where the textbook is found to be inadequate.

Keller and Sherman (1974) indicate that a well-written procedure section tells the students what to do, how to self-test comprehension, how to decide whether to proceed or review and how to decide when they have finished. The study guide may include any further information that helps the students to complete their work successfully, the main criterion for a study guide being that it works.

Student Proctors

These are usually students who have successfully completed the course. Their job is to score quizzes, discuss answers, tutor students who are experiencing difficulties, and provide feedback to the instructor on the course in general.

According to Keller, the proctor is a vital component of the system, being a primary source of external reward for the mastery of each unit; they constitute the personal contact between student and system and provide immediate feedback to the students. The proctor allows interaction between the student and the system and a degree of flexibility that can only come from human interaction: credit can be given to a student if he appears to understand the item tested, even though the answer to the item was incorrect due to some minor error. Also the proctor can discuss correct responses in order to determine whether the answer is guessed or arrived at for invalid reasons. This is what Keller says of the 'proctor':

> The proctor is the new link in the chain. His principal job is that of decreasing the gap of understanding between the student and teacher - a gap that is sometimes very wide. He can do this because he has a repertoire of behaviour that is intermediate between the two with respect to the subject matter of the course. He may find logic in an answer that an instructor would never have perceived; he may restate a question in a way that brings out evidence of knowledge; and he may by example or parallel statement strengthen a student's grasp of a concept. He provides the individual consideration that a student may never have had before. (Keller and Sherman, 1974, p.20-1)

The proctor is helped in his work by 4 factors:

1. his experience with a similar course in which he was recently successful;
2. a weekly proctor's meeting in which every question on every test may be discussed;
3. a handy list of acceptable answers to the questions, to jog his memory if needed;
4. usually, relevant knowledge from advanced courses within the same subject-matter field.

Those proctors, described above, are called by Keller 'grading proctors', and he estimates that ten or more would be required for a class of one hundred students. Two further proctors may be involved in the general scheme: (1) the testing proctor; and (2) the study hall proctor.

1. Testing proctor: is responsible for all record- keeping and actually administering the tests. Keller says that this non-educational role may be undertaken by a reliable clerk or by grading proctors on a rotational basis.

2. Study hall proctor: this position is so important that it may be filled by the instructor or his assistant. He stands between the textbook and the student (just as the grading proctor stands between student and the instructor). He clarifies obscure passages, helps with difficult operations, explains unintelligible references etc. The study hall proctor generally serves a small percentage of the class (usually its weakest members) and the number of cries for help will decrease as study habits are improved and students gain in self-reliance.

Keller says that the study-hall proctor may be selected on the basis of course history, achievement or special aptitudes but also cautions that if he acts too much like a teacher he may not be the most effective promoter of learning in the classroom.

In addition to the proctors there is the classroom assistant whose tasks include distribution of assignments, study questions, checking of supplies, maintenance of progress charts etc. Later, he may take the place of the instructor at various meetings and in general give the instructor time to alter the course as and when necessary.

The instructor's role is that of a manager of learning. He selects and analyses the material to be mastered, he decides how to present it, and constructs the various questions based upon it. Keller emphasizes that the textbooks must be carefully read before selection to determine their suitability for students who will be questioned upon it in detail - errors, confusions and contradictions that might go unnoticed in a conventional system will stand out as 'unsightly blemishes' in PSI.

The instructor is also responsible for dividing material into suitable units, devising suitable study questions to help students and test questions for the mastery test. He will receive a great deal of information on everything he has devised for the course and as a result he will be responsible for the continual modification of the system, as he aims for perfection.

Lectures

Lectures are used for enrichment rather than to transmit information, which in PSI is done through the use of study guides and the text. Lectures are given to demonstrate the excitement of the discipline; and may be used to show films or for other special events such as visiting speakers. They serve a purely supplemental function. As Keller and Sherman point out:

> A few lectures are useful. The instructor can serve as a model, impart style, even provide inspiration. In fact we may expect a very polished and challenging presentation now that such events are infrequent. Such lectures must be fine indeed; otherwise they will not attract an audience. Since the lecture is given for motivational reasons rather than for the transmittal of essential information, PSI lectures are 'extras' and not required. (Keller and Sherman, 1954, pp.41-42)

Personalized System of Instruction: The Research

In the 'Review of Research in Education' (1975) McKeachie & Kulik summarized the results of the available evaluations of PSI:

1. Keller plan is an attractive teaching method to most students. In published reports students rate Keller plan more favourably than teaching by lecture.
2. Self-pacing and interaction with tutors seem to be features most favoured by students. There are reports of higher-than-average student drop-out rates for Keller courses.
3. Content learning, as measured by final examinations, is adequate: in published studies final examination performance always equals and usually exceeds performance in lecture sections.
4. Students almost invariably report that they learn more from PSI and report expending more time and effort.

In addition to these generally favourable reports evidence was also accumulating which showed the relative importance of the various components of PSI in contributing to the system's overall successful performance.

Student Ratings of the Course

Gallup (1970) found favourable comments on a psychology PSI course from 98% of students and Born & Herbert (1971) found that only 7 of 145 students would not recommend the course to other students. When looking at Library Science, Knightly & Sayre (1972) found that the self-paced mode of instruction was considered better than lecture courses by 100% of students; the course being rated as 'above average' or 'one of the best' by 93%.

Perhaps it should be noted here that in the early days many academics tended to describe PSI as a de-humanizing, programmed approach to instruction, and stress its links with Skinner's behaviourist orientation. Students

who studied such courses, however, liked its freedom, self-determination and personal interaction!

Features Rated Highly by Students

Green (1971) categorized student comments about Keller courses and reported that the feature most often praised is the freedom of pace. In another study, Nelson & Scott (1972) had students rank 12 features of PSI. Most important were (1) self-pacing; (2) interactions with teaching assistant or instructor; (3) small unit steps; (4) discussions with proctors. Self-pacing was also seen to be the feature rated most highly in a similar study by Myers (1970). Clearly, for the students, the most attractive features of PSI are self-pacing and personalized interactions.

Green (1971) reported that 90% of students on Keller plan took further Keller options when offered; and Gallup (1974) reported that he had to limit student numbers on his course when it was introduced as a Keller type option.

This overall positive response to PSI was not consonant with McKeachie & Kulik's (1975) suggestion that student drop-out was higher than the average for conventional lecture courses. For example, Keller presented data that showed withdrawal rates of about 20% - no comparison data was available, but 20% does seem to be very high. Sheppard & MacDermot (1970) reported a drop-out rate of 17% for a Keller course on introductory course, compared with 6% for a traditional class. Born (1971) also presented data showing a withdrawal rate of 14%, 25% and 14% for Keller sections compared with 5%, 6% and 5% for conventional courses.

However, it is not the whole story. McMichael & Corey's (1969) results showed a high drop-out rate for the PSI course, but a slightly higher one for the conventional lecture course: 12% compared with 13%. Kulik, Kulik and Cohen (1979) in their meta-analysis of PSI research found that of 27 investigations which reported comparisons of course completion, the differences between PSI and conventional courses were trivial.

Procrastination did seem to be a major disadvantage of such a self-paced system of instruction in the early days and Green (1971) and others suggested ways in which it could be prevented:

1. publishing a recommended schedule of dates for passing units, which will yield a steady rate of work,
2. making admittance to 'fun' lectures contingent upon passing units,
3. offering early final exams for students finishing the course early.

Content Learning

Kulik, Kulik and Smith (1976) presented details of 39 studies which satisfied their basic requirements of sound experimental design. In 38 of the 39 studies final exam performance was better in the PSI course. In 34 of the comparisons the difference was considered to be statistically reliable. The authors state that the differences were large enough to have practical significance: they show that for a typical result a student, who is average on a standard test, scoring 50% in a typical lectures-based course, would score 75% if she takes the PSI courses, indicating an Effect Size of 0.66. A later meta-analysis (Kulik, Kulik, and Cohen, 1979) found 61 studies which compared PSI and conventional teaching. PSI final examinations were 8 percentage points higher, giving an average effect size of 0.49. Also, there was 35% less variation in the scores for the PSI group. These results refer to final examinations. But what about the long-term impact PSI has on students? Again, Kulik, Kulik and Cohen (1979) indicated that PSI comes out with better results. In fact the differences were greater after the retention period (which varied from 3 weeks to 15 months), suggesting that final exams, if anything, underestimate the magnitude of PSI's effect.

Another way in which the effects of PSI can be investigated is the transfer study, where students study, for example, MATHS1 by either PSI or conventional methods. They then study MATHS 2 together in a conventional class. How does PSI affect performance in MATHS 2 ? The transfer studies all indicate that those students who take the PSI option for the initial course out-perform conventionally taught students on the second course, even though the second course is a conventionally taught course.

Component Analysis

There can be no doubt that PSI is effective, the question is, why is PSI so effective? Various studies have attempted to throw some light on this question, and have looked at the various components of PSI.

The research studies tend to show that small units are important to the effectiveness of PSI - if the units become too long PSI's effectiveness deteriorates (Calhoun, 1976). Semb (1974) also found better performance in major exams for students working on smaller units. Nelson and Bennett (1973) and O'Neill et al (1975) found more consistent progress in shorter unit PSI courses.

Feedback also seems to be important and there are several studies in which groups receiving delayed feedback were compared with groups receiving immediate feedback (Calhoun, 1976; Farmer et al, 1972; and Johnson and Sulzer-Azaroff, 1974). The achievement of the delayed feedback groups was significantly lower than groups receiving immediate feedback.

Relaxation of the mastery criterion leads to a decline in the effectiveness of PSI, as was anticipated by Keller. Semb in 1974 showed that a high mastery criterion (100%) produced better final test performance than a low mastery criterion (60%) and this was confirmed in a later study (Phillips and Semb, 1975).

Surprisingly, the proctors don't seem to be an essential feature of PSI. Blackburn, Semb and Hopkins (1975) compared a Keller PSI course with a section which included self-grading (the proctors only acting as admin-

istrators). In the final exam both groups performed equally well. Calhoun (1976) compared written feedback with feedback provided by the proctor and again there were differences, but they were not statistically significant.

Research studies have also failed to show that self-pacing is important. Several studies show no difference on final examination or student evaluation for teacher-paced and self-paced groups (Beneke and Taylor, 1975; Calhoun, 1976; Lewis, 1972).

Although Keller recommends lectures in PSI courses as motivational agents, some instructors have built in conventional lectures in addition to the individualised work. Calhoun (1976) found that twice-weekly lectures did not improve the performance of PSI students and Minke and Carlson (1973) found that adding lectures did not change the percentage of students passing, the withdrawal rate or the overall evaluation of students passing the course.

LFM AND PSI: A REVIEW OF THE RESEARCH

The research on mastery learning as found in the approaches of Bloom and Keller was summarized in a comprehensive review by Block and Burns (1977). In this paper the authors reported a meta-analysis on the research, giving the effect size for LFM achievement comparisons and a combined retention effect size for PSI and LFM.

A total of 97 comparisons were identified, involving different types and numbers of students. Mastery students scored higher than non-mastery students 89% of the time, and statistically significantly higher 61% of the time. For retention (combining PSI and LFM) the mastery students were almost always superior to the non-mastery groups, significantly so 63% of the time. The average effect size for the LFM-taught students was 0.83. A similar large effect size was obtained from studies of LFM and PSI which included a retention test, with an average retention effect size of 0.67.

In addition, Block and Burns looked at the variability in learning. According to Bloom's theory the mastery approach will lead to greater uniformity in learning when compared to traditional methods, which should show increasing variability. In 80 studies comparing variance in achievement test scores, mastery students showed less variation in their scores 74% of the time, and less variability in retention scores 85% of the time. The reduction in variance was approximately 50% for both achievement and retention measures.

Kinds of Learning

Block and Burns also looked at learning in qualitative terms. They found some evidence to suggest that mastery strategies may be more likely to elicit higher order than lower order behaviours. Several researches found that mastery students were superior on essay-type questions that included comprehension, application, analysis, synthesis and evaluation skills, and that mastery students were more able to deal with these complex high order skills than the lower order multiple-choice knowledge questions.

There was some evidence, however, that for retention periods of 6 or 18 months it was the lower order behaviours associated with knowledge, rather than the complex application or analysis-synthesis, that were retained at a higher level in the mastery group.

The mastery approaches typically elicited more favourable affective responses from students, in particular there was a positive impact on students' interest in and attitudes to the subject matter, and also on self-concept, academic self-confidence, attitudes to co-operative learning and towards instruction. Although there is some evidence to suggest that mastery set at the 100% level adversely affects such attitudes and that an optimum figure of 90% maintains both the positive affective and achievement responses.

Time Consequences

A group of studies investigated the amount of time spent in learning in mastery and non-mastery groups. Uniformly, they found that mastery strategies had a homogenizing effect on differences in study time. One study found that differences were reduced from 7 to 1, to 4 to 1 by mastery methods, and that there was progressively less variability in the mastery study time. In the final unit of eight there was 90% less variability.

Block and Burns comment that this reduction in individual differences in study time is purchased at the price of additional study time for the slower students. Over the short term there is an increase in the average study time required for mastery of 70-80%, compared with non-mastery. In the long term this may be reduced to 20-50%.

The extra time is available from two sources. The first is the teacher who is encouraged to prepare and organize the instructional plans, procedures and materials outside of class time and prior to the instruction. This allows more time for interaction with the students. The student can also be encouraged to exert more effort, especially in overcoming procrastination, and various incentive schemes have been used, such as frequent testing, varying the size of learning units, teacher/peer pacing of instruction. The emphasis is clearly on techniques that encourage teachers and students to spend more time learning as opposed to non-learning activities.

Achievements of the Behavioural Systems

In many ways the work of Bloom and Keller in their mastery learning systems represents the summit of achievement for the neo-behaviourist 'systems' approach to instruction. Both were built on firm, though differing, psychological principles and of all the applications of the behavioural science approach to the technology of education, LFM and PSI represent the only methods which consistently produce significant educational results. Clark (1983) attributes this to the emphasis placed on method, defined by Glaser (1976) as 'the conditions which can be implemented to foster acquisition of competence,' rather

than on the media involved in the delivery of instruction, as represented by the physical science approach.

It is fitting that Benjamin S. Bloom, whose work has been central to the development of instructional science for a generation, should have the last words:

> If you can be moved to try these ideas with a few teachers in your school for even as short a period as three months, you can determine the validity and limits of these ideas where they belong - in your classrooms and with your teachers and students.

He continues:

> However, neither further opportunity for education nor increased financial support for education will do much to improve the education of each of our students. The answer does not lie in additional funds, new fads, or major and sweeping changes in the organization of our educational system. As I see it, the solution lies in our views about students and their learning. These have grown out of our practices and they will not be changed until we alter these practices. When the changed practices succeed in promoting more effective learning, both teachers and students will change their views. (Bloom, 1978, p.563)

REFERENCES

Beneke, W.M. and Taylor, D.C. (1975) *Effects of Unit Deadlines on Performances in a PSI Course*. Paper presented at the annual meeting of the American Psychological Association, Chicago.

Bergan, J.R. (1980) 'The structural analysis of behaviour: an alternative to the learning-hierarchy model', Review of Educational Research, 50, 4, 625-646.

Blackburn, T., Semb, G. and Hopkins, B.L. (1975) 'The comparative effects of self-grading versus proctor grading on class efficiency and student performance'. In J.M. Johnson (Ed) *Behaviour Research and Technology in Higher Education*. Springfield Ill.: Charles C. Thomas.

Block, J.H. (1971) *Mastery Learning*. London: Holt Rinehart & Winston.

Block, J.H. (1974) *Schools, Society and Mastery Learning*. London: Holt, Rinehart and Winston.

Block, J.H. and Burns, R.B. (1977) 'Mastery learning'. In L.S. Shulman (Ed) *Review of Research in Education. Volume 4*. Itasca, Ill.: F.E. Peacock, 3-49.

Bloom, B.S. (1964) *Stability and Change in Human Characteristics* . New York: John Wiley.

Bloom, B.S. (1968/81) 'Learning For Mastery', (UCLA-CSEIP) *The Evaluation Comment*, 1(2). In B.S. Bloom (Ed) *All Our Children Learning*. London: McGraw-Hill.

Bloom, B.S. (1976) *Human Characteristics and School Learning.* New York: McGraw-Hill.

Bloom, B.S. (1978) 'New views of the learner - implications for instruction and curriculum', *Educational Leadership*, 35, 562-76.

Bloom, B.S. et al. (1956) *Taxonomy of Educational Objectives Handbook I: Cognitive Domain.* New York: McKay.

Bloom, B.S., Hastings, J.T. and Madaus, G.F. (1971) *Handbook on Formative and Summative Evaluation*. London: McGraw-Hill.

Born, D.G. (1971) *Student Withdrawals in Personalized Instruction Courses*. Paper presented at the meeting of the Rocky Mountain Psychological Association, Denver, May.

Born, D.G. and Herbert, E. (1971) 'A further study of personalized instruction in large university classes', *Journal of Experimental Education*, 40, 6-11.

Bruner, J.S. (1966) *Toward a Theory of Instruction*. Cambridge, Mass.: Harvard University Press.

Calhoun, J.F. (1976) 'The combination of elements in the personalized system of instruction', *Teaching of Psychology*, 3, 73-76.

Carroll, J. (1963) 'A model of school learning', *Teachers College Record*, 64, 723-733 .

Clark, R.E. (1983) 'Reconsidering the research on learning from media', *Review of Educational Research*, 53, 4, 445- 459.

Dubin, R. and Taveggia, T.C. (1968) *The Teaching Learning Paradox.* University of Oregon Press: Eugene.

Farmer, J., Lachter, G.D., Blaustein, J.L. and Brett, K. (1972) 'The role of proctoring in personalized instruction', *Journal of Applied Behaviour Analysis*, 5, 401-404.

Gagné, R.M. (1962a) 'The acquisition of knowledge', *Psychological Review*, 69(4), 355-365.

Gagné, R.M. (1965) *The Conditions of Learning (1st Edition).* London: Holt, Rinehart and Winston.

Gagné, R.M., Mayor, J.R., Garstens, H.L. and Paradise, N.E. (1962) 'Factors in acquiring knowledge of a mathematical task', *Psychological Monographs*, 76(7), No.526.

Gagné, R.M. and Paradise, N.E. (1961) 'Abilities and learning sets in knowledge acquisition', *Psychological Monographs*, 75(14), Whole No. 518.

Gallup, H.F. (1970) *Individual Instruction in an Introductory Psychology Course.* Paper presented at the meeting of the American Psychological Association, Washington, D.C. (Sept.)

Gallup, H.F. (1974) Problems in the Implementation of a course in Personalized Instruction. In J.G. Sherman (Ed) *PSI: 41 Germinal Papers*. Menlo Pk., Cal.: Benjamin.

Glaser, R. (1976) 'Components of a psychology of instruction: towards a science of design', *Review of Education al Research*, 46(1), 1-24.

Green, B.A. (1971) 'Physics teaching by Keller Plan at MIT', *American Journal of Physics*, 39, 764-775.

Johnson, K.R. and Sulzer-Azaroff, B. (1974) *The Effects*

of Different Proctoring Systems on Student Examination Performance and Preference. Paper presented at Second National Conference on Research and Technology in College and University Teaching. Georgia State University, Atlanta, GA. (Nov.)

Keller, F.S. (1968) 'Goodbye, Teacher', *Journal of Applied Behaviour Analysis*, 1, 79-87.

Keller, F.S. and Sherman, J.G. (1974) *The Keller Plan Handbook*. Menlo Park-California: Wa. Benjamin.

Kim, H. (1971) 'Mastery learning in the Korean middle schools', *UNESCO Regional Office for Education in Asia*, 6(1).

Knightly, J.J. and Sayre, J.L. (1972) 'Self-paced instruction for library science students', *Journal of Education for Librarianship*, 12, 193-197.

Kolb, J.R. (1967-68) 'Effects of relating mathematics to science instruction on the acquisition of quantitative science behaviours', *Journal of Research in Science Teaching*, 5, 174-182.

Kulik, J.A. and Jaksa, P. (1977) 'PSI and other educational technologies in college teaching', *Educational Technology*, 17, Sept. 12-19.

Kulik, J.A., Kulik, C.C. and Cohen, P.A. (1979) 'A meta-analysis of outcome studies of Keller's personalized system of instruction,' *American Psychologist*, 34(4), 307-318.

Kulik, J.A., Kulik, C.C. and Smith, B. (1976) 'Research on the personalized system of instruction', *Programmed Learning and Educational Technology*, 13, 21-30.

Lee, Y.D. et al. (1971) *Interaction Improvement Studies on the Mastery Learning Project - Final Report on Mastery Learning Programme*. Educational Research Centre, Seoul National University. (Nov.)

Lewis, J.D. (1972) *Manipulation of Two Variables in a Contingency-managed Statistics Class*. Paper presented at the West Coast PSI Conference. California State University.

McKeachie, W.J. and Kulik, J.A. (1975) Effective college teaching. In F.N. Kerlinger (Ed) *Review of Research in Education Vol.3*. Itasca, Illinois: F.E. Peacock.

McMichael, J. and Corey, J.R. (1969) 'Contingency management in an introductory psychology course produces better learning', *Journal of Applied Behaviour Analysis*, 2, 79-83.

Minke, K.A. and Carlson, J.G. (1973) *A Comparison of Two Lecture Systems in Unit Mastery Instruction.* Paper presented at the annual meeting of the American Psychological Association, Montreal. (Aug.)

Morrison, H.C. (1926) *The Practice of Teaching in a Secondary School*. University of Chicago Press: Chicago.

Myers, W.A. (1970) 'Operant learning principles applied to teaching introductory statistics', *Journal of Applied Behaviour Analysis*, 3, 213-220.

Nelson, T.F. and Bennet, M. (1973) 'Unit size and progress rates in self-paced instruction', *Journal of College Science Teaching*, 3(2).

Nelson, T.F. and Scott, D.W. (1972) 'Personalized instruction in educational psychology', *Michigan Academician*, 4, 293-302.

O'Neill, G.W., Johnston, J.M., Walters, W.M. and Rasheed, J.A. (1975) *The Effects of Quantity of Assigned Material On College Student Academic Performance and Study Behaviour*. In J.M. Johnston (Ed) *Behaviour Research and Technology in Higher Education*. Springfield, Illinois: Charles C. Thomas.

Phillips, T.W. and Semb, G. (1975) *A Comparison of Pre-quiz and Post-quiz Lectures in a Contingency-managed University Course.* Paper presented at the annual meeting of the American Association, Chicago. (Sept.)

Rowntree, D. (1977) *Assessing Students: How Shall We Know Them?* London: Harper and Row.

Ryan, D.W. and Schmidt, M. (1979) *Mastery Learning: Theory, Research, and Implementation*. Toronto, Ontario: The Ontario Institute for Studies in Education.

Semb, G. (1974) 'Personalized instruction: the effects of mastery criteria and assignment length on college student test performance', *Journal of Applied Behaviour Analysis*, 7, 61-69.

Sheppard, W.C. and MacDermot, H.G. (1970) 'Design and evaluation of a programmed course in introductory psychology', *Journal of Applied Behavioural Analysis*, 3, 5-11.

Skinner, B.F. (1954) 'The science of learning and the art of teaching', *Harvard Educational Review*, 24(2), 86-97.

Tyler, R.W. (1932) 'The construction of examinations in botany and zoology', *Service Studies in Higher Education*. Ohio State University, Bureau of Educational Research Monographs, 15, 49-50.

Washburne, C.W. (1922) 'Educational measurements as a key to individualizing instruction and promotions', *Journal of Educational Research*, 5, 195-206.

White, R.T. (1973) 'Research into learning hierarchies', *Review of Educational Research*, 43(3), 361-375.