

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

THE EFFECTS OF TEACHER ATTENDING  
AND RESPONDING BEHAVIOURS ON  
PUPIL ACHIEVEMENT.

A thesis presented in partial fulfilment of the requirements  
for the degree of Master of Philosophy in Education at  
Massey University.

*Stephen Frederick Steffens*

1978

## ABSTRACT

This thesis reports a field experiment of teachers' non-verbal attending and verbal responding behaviours and their relation to pupil achievement. It was expected that teachers who used higher level non-verbal and verbal skills would have children who achieved at a higher level.

A group of fifteen teachers, five in three different schools, were video-taped at the beginning of the school year. Because all of the teachers were functioning at low levels on scales of attending and responding, one group of teachers was given ten hours of training in these skills. One other group of five teachers was given a placebo treatment and the remaining group received no treatment. All of these teachers were video-taped at the end of the school year.

All of the children were tested at the beginning and the end of the school year on standardized achievement measures.

Analysis of the data revealed that ten hours of training given one hour at a time at intervals of one week was not sufficient to establish average classroom conditions that were significantly different from the other experimental treatments. The trained teachers did improve in the expected direction more than the other teachers.

The students of the two teachers who responded at the highest levels were compared to the students of the two teachers who responded at the lowest level. The results supported the hypothesis.

From this sample of teachers those who were female, who responded to feelings, who had students from a higher socio-economic status, who were just beginning to teach, and who attended and responded appropriately had students with higher results.

## ACKNOWLEDGEMENTS

Any such undertaking as this that directly involves fifteen teachers and more than four hundred primary and intermediate pupils can only be successful with cooperation and support from many individuals. To all of these many individuals I wish to give my sincere appreciation:

the three headmasters, fifteen teachers, and more than four hundred pupils for their efforts, curiosity, and openness. I am only sorry that some of these outstanding people must here remain anonymous,

Selwyn Cathcart for technical brilliance and good natured tolerance and alertness to one who threatened to dismantle expensive equipment unknowingly,

Professor C.G.N. Hill and several of the Education Faculty who gave support to this project in many ways,

Selwyn Moore, Senior Inspector for the Wellington Education District, who gave official approval and showed personal interest in the progress of this project,

Marriane Wheatley and Graham Woolford for long hours viewing video-tapes and making hundreds of decisions while they coded, Cathy Smith and Nola Simpson of the Computer Unit who ironed out bugs in version six of S.P.S.S. and then searched with me to find what was not a bug,

Jennifer Darby for reading a draft copy and smoothing out communication links with the reader,

Hazel Hayward for transforming the manuscript into a final, very presentable unit.

To these following few I have special gratitude:

Dr. A.C. Webster who lives by and as a model, available at many levels to respond to the needs felt and unfelt, known and only faintly imagined, with respect and care, with an excitement and optimism that good can be done,

Lis, Kevin, and Erik who have reminded me frequently of the valuable priorities available in life and encouraging us to live what is important.

## TABLE OF CONTENTS

	<u>Page</u>
ACKNOWLEDGEMENTS	iv
LIST OF TABLES	vi
CHAPTER I THE PROBLEM AND ITS ANTECEDENTS	1
Education for Learning	1
Motivation and Learning	2
Self-Concept and Performance	2
Weaknesses of Research Linking Self-Concept with Achievement	6
Carkhuff and the Operationalization of Helping	7
CHAPTER II THE HYPOTHESES	10
CHAPTER III RESEARCH DESIGN	13
CHAPTER IV INSTRUMENTATION	17
Scale for Attending	17
Scale for Teacher Responding and Initiating	20
Measures of Pupil Achievement	26
Measure of Pupil Attitude	27
CHAPTER V TRAINING PROCEDURES	29
Experimental School	29
Placebo School	31
CHAPTER VI PROCEDURE	42
CHAPTER VII DATA ANALYSIS	47
Attending - Reliability	47
Teacher Attending Behaviour	49
Responding - Reliability	52
Teacher Responding Behaviour	53
CHAPTER VIII DISCUSSION AND CONCLUSIONS	67
BIBLIOGRAPHY	72

## LIST OF TABLES

<u>Table</u>	<u>Page</u>
1 The Research Design	14
2 Pupil-Teacher Communication	17
3 Scale for Teacher Attending	19
4 An Adaptation of the Carkhuff Scale for Effective Helping for use in Classroom Observation	22
5 The Teacher-Pupil Question Inventory	34
6 The Florida Taxonomy of Cognitive Behaviour	37
7 Product Moment Coefficients of Teacher Attending Scores	48
8 Analysis of Variance: Teachers Attending Pre-test	49
9 Analysis of Variance: Teachers Attending Post-test	50
10 Increases in Attending by Teacher Sex	51
11 Analysis of Variance: Teacher Responding Pre-test	53
12 Analysis of Variance: Teacher Responding Post-test	54
13 Average Gain in Responding	55
14 Lower Primary Teachers' Average Scores	56
15 Analysis of Covariance: High Teacher vs. Low Teacher - Reading Comprehension	57
16 Analysis of Covariance: High Teacher vs. Low Teacher - Vocabulary	57
17 Analysis of Covariance: High Teacher vs. Low Teacher - Mathematics	58
18 Correlation Matrix	60
19 Multiple Regression onto Average Student Gain	61
20 Rank Ordering of Teachers According to Average Pupil Gain	63
21 Students' Parents' Socio-Economic Status	64
22 Pupil Attitude of Teacher	66

## THE PROBLEM AND ITS ANTECEDENTS

### Education for Learning

Education for cognitive development has been the major emphasis in schools. However, there is increasing dissatisfaction, not with the aim to produce more knowledgeable people, but with learners who leave school with inadequate cognitive skills. The plan and purpose of this thesis is to unravel some of the contributions of teachers to the success or failure of their pupils at meeting some of the aims of education.

One basic supposition is that all learning begins with the learner. The motivation of the learner is his basic energy available and directed toward achieving his goal. Teachers who can utilize this energy will achieve more. How can a teacher know the motivation of a student? How can a teacher direct the motivation of a student?

One popular approach to these questions has arisen out of self-theory. The investigations and logic of self-theorists will be presented along with their conceptual and methodological shortcomings.

Another approach to the problem of motivation that will be examined is best exemplified by Robert Carkhuff's research in counselling. Certain observable behaviours have been identified as crucial to the success of counselling. These skills will be presented and adapted for use in the classroom and tested for their effectiveness.

Carl Rogers (1969) in a Harvard Conference on Classroom Approaches to Influencing Behaviour irritated many when he said, "The outcomes of teaching are either unimportant or hurtful," ... "Do away with teaching," ... "Do away with examinations as they measure only the inconsequential type of learning," as well as other proposals to

facilitate learning that is self-discovered and self-appropriated. Rogers' basic assumption was that children and all adults are initially and constantly motivated. In his view teachers "do not have to turn kids on."

### Motivation and Learning

Teachers have been greatly influenced by the conceptions of motivation that educational psychologists have advanced. Originally, motivation was thought to be a drive state that would arise to meet a bodily need. When the need was satiated, the motivation would no longer exist and activity would theoretically cease. Some psychologists inferred needs from the activity of the organism. Murray (1938) was able to infer twenty-seven needs to account for the human activity that he observed. Of course the way was left open to infer more needs if the twenty-seven were not adequate. It was this growing number of motivational drive states and the anomalies of life forms engaging in activity that could not be described as meeting the needs aroused by a drive state, that aroused the psychologist's need for parsimony. Would still more motivational needs have to be postulated for infants that played, monkeys that worked for ten hours manipulating a puzzle, (Hunt, 1968) or mice that preferred an environment that had stripes on the wall to one that had no stripes?

Hebb (1955) pointed out that the multiplication of drives would not solve the problem of motivation. Even at the level of individual nerve cells there is activity at a rather constant background rate. With outside stimulation, this rate changes. But by itself, the nerve cell and also the entire organism is by nature active.

### Self-Concept and Performance

It is a large jump from the level of analysis focused on individual nerve cells to the level of classroom interaction. However, it is necessary to observe that an individual's activation level affects his classroom behaviour. The responsibility of a teacher is therefore broadened by having self-activating pupils. (Day and Berlyne, 1971) Still necessary is the traditional teacher's role of being a stimulus field for the learner. Another role is to know the motivational level of the learner. Those educators who emphasize the importance of the

learner's self-concept are attempting to measure the individual's level of motivation. How can a teacher best acquire the necessary skills involved in being able to know the motivational level of each learner and therefore teach each learner most effectively?

Jersild (1955) suggests that the teacher should first know himself and then, by being open, begin to know the pupil as a person. This knowing of the self, both by the teacher and the student, will lead to an increased ability to "think straight" as feelings often influence the cognitive flow of information. (Brown, 1971) By relating the self to the subject matter, the content can acquire personal relevance and a greater depth of involvement will result. Of course, the obverse is also possible. That which society has deemed important, an individual may think is irrelevant and therefore decide to forego the acquisition of that section of the curriculum.

Jersild (1951) also believes that the child is more finely equipped to deal with learning than is assumed by most teachers.

"Human beings, from an early age, have more capacity for learning to face and to understand and to deal constructively with the realities of life than we have hitherto assumed in our psychological theories or in our educational practices." (p. 122)

The close paralleling of a person's self-concept and his behaviour has frequently been noted.

"(It is) ... not the people who feel that they are liked and wanted and acceptable and able who fill our jails and mental hospitals. Rather, it is those who feel themselves deeply inadequate, unliked, unwanted, unacceptable, unable, and the like." (Combs, 1965, p. 563)

Combs further suggests that teachers should include the self of each individual as part of the curriculum. His concern with knowing and improving the student's self-concept is summarized with this question: "Does this (any part of the curricula or any method) help our students to feel more liked, wanted, accepted, able, dignified, worthy, important, and so on?" If a person has this empathic ability to recognize the self of each student and to foresee the effects of his

actions on the self of each student, Combs believes that this person can do well at teaching with no training. Whereas, without this skill, even with teacher training as it usually is at present, the individual may still do poorly.

An increasing number of researchers have been led to investigate the initial determinants of the self-concept of the child, the results in life as a function of the self-concept, and the re-shaping or enhancement of a person's self-concept.

Swift (1971) in England found that children from the middle class had a much greater chance for selection for higher education than children from the working class. Social class alone as a predictor has shortcomings as it neglects the internal dynamics of the family, for instance, the sons of working class fathers who were dissatisfied with their occupation did well on the national selection examination.

Harrison (1968) admitted that generally the social class background of an individual was a good predictor of a large constellation of attitudes, behaviours, and environmental conditions that affect the individual's self-concept and his academic achievement. However, there were anomalies of advantaged students doing poorly academically and students with disadvantaged backgrounds doing well academically. Harrison found that the attitudes of the achievers, irregardless of their social class, were similar. The self-concept of a person was more of a major predictor of performance in school than his social class.

The teaching behaviour of mothers of high and low social class was observed while they were teaching their own pre-school children. (Hess and Shipman, 1971) There were significant differences in the amount of imperatives used, the amount of language used, the concern for the future, and the relationship of rationality and authority. The high social class children were developing self-concepts and skills that gave them confidence to act on their own initiative with rational behaviour.

Carl Rogers (Rogers, et al, 1948) studied the predictive power of eight factors on the improvement of delinquent adolescents. The most powerful predictor was the level of self-understanding followed in order by: social competency, hereditary background, intelligence, family atmosphere, economic level of the family, physical health, and amount of education. Rogers proposed that to change a person for the better, it would be necessary to change the self-concept by aiming

toward more self-understanding which leads to attitudes of self-control and self-direction.

While these results are certainly unexpected and stimulating, they should be treated with slightly more caution than was exhibited, especially when it is known that this research was the first ever reported of an observer inferring levels of personal insight from case history reports. (Wylie, 1974) The statistical analysis of the data was not complete enough to justify the strong conclusions. The simple correlational findings, while quite high, may have had other underlying common factors that could have been extracted by the techniques of multiple regression or even partial correlation. More will be said below about the weaknesses of this research and other research in self-theory that attempt to determine the motivational component in student learning. First, some of the research in self-concept as it relates to performance in schools will be recounted.

The importance of the self-concept in relation to school achievement became apparent to Fink (1965) in his study of underachievement. His research gave evidence of ten environmental variables that were predictors of student achievement. He was dissatisfied with this multiplicity and postulated that the environmental influences affected the self-concept, for better or worse, and that the resultant self-concept was the vital determiner of achievement.

Davidson and Lang (1960) reported that Reeder observing in the primary school and Stevens observing in the tertiary level both found positive self-feelings positively correlated with higher academic achievement. They postulated that the teacher's approach can initiate a self-perpetuating circle of, "The teacher thinks I am good. I think I am good. I behave appropriately. I achieve high results. The teacher thinks I am good."

Brookover (1964) concluded that a pupil's self-concept is significantly related to performance when the ability factor is controlled statistically. The self-concept is influenced by those people who are important to the child. He suggests that if it is possible to enhance the self-concept, then the person's performance will be raised.

Staines (1958) conducted a small experiment using two teachers who were judged to be effective by their superiors and by their past records. By removing the stress on performance and not ranking students in the one class, Staines successfully helped one teacher to improve the self-concept of the children in her class. The students set

personal learning goals and carried them out. This class had higher gain scores on an academic achievement measure and also exhibited an improved self-concept. The question of causation was left open from this experiment. Was the self-concept higher in the experimental class because of their improved performance or was the improved performance a result of an improved self-concept? Part of the cause for the improvement of both may be the result of removing stress in the children and accepting each pupil. Such strategies of improving self-concept and performance are becoming apparent.

In the summarization of some of the key findings of the observation of classroom behaviour it was found that warmer and more enthusiastic teachers produced better achievement gains and also more positive atmospheres in their classrooms. (Brophy and Good, 1974)

From working with adults, frequently in the area of business management, McClelland (1965) advances the proposition that an individual's self-concept will influence his behaviour. If it is advantageous to improve another person's self-concept, it is more likely to improve in an interpersonal atmosphere where the individual is warmly, but honestly supported and respected.

#### Weaknesses of Research Linking Self-Concept and Achievement

The study of the predictive power of the self-concept done by Rogers (Rogers, et al, 1948) is indicative of several weaknesses of analysis and conceptualization that is frequent in research relating to self-concept and performance. The scales employed for measuring the different characteristics were treated in the statistical analysis as interval scales. This assumption was not justified. The reliability of the ratings was not measured for this analysis. For the same scales in a different experiment the reliability was low. On a seven step scale only four judges out of six would be within two scale steps of each other. (Wylie, 1974)

A much greater weakness was the assumption that the high correlation that was found between the amount of insight and eventual outcome indicated a causative factor. This assumption is quite frequent in the literature of self-concept research.

"The demonstrated correlation between the student's self-concept and the evaluations that the student perceives others to hold of him leads to the question as to whether it is possible to change the self-concept. Furthermore, if self-concept is in fact a key factor in role performance, changes in self-concept should result in changes in performance." (Brookover, et al, 1964)

The recognition that success builds positive self-concepts and that success is dependent upon the possession of necessary skills is a more fruitful line of inquiry. This approach could point to an earlier common determiner of both the individual's self-concept and his performance ability. It is easy to obtain measures of self-concept but as Wylie (1974) points out, there are many pitfalls in the pursuit of validity of such measures.

One fruitful conception is to consider self-concept as a process and even an outcome factor and not as a determiner of desired outcomes. Parents who train their children with acceptable skills give positive reinforcement when desired skills are carried out. These reinforcers are seen by the child as validations for his activity and are generalized to positive self-concepts. In this manner the skills and a positive self-concept are linked. A re-training programme for pupils with low self-concept and inadequate delivery skills will be ineffectual if it concentrates solely on improving the self-concept in vacuo. Programmes that aim toward changing behavioural skills usually reinforce the execution of these skills which leads to improved delivery and an improved self-concept.

#### Carkhuff and the Operationalization of Helping

The combination of theory and positive research findings for improving a person's self-concept and achieving success comes from Robert Carkhuff (see especially Carkhuff, 1969) and those who have utilized his achievements and built on them. Working initially out of a counselling framework, he found that the more accurately the therapist could understand the client's viewpoint and share his understanding with the client, the more helpful the relationship would be. He utilized the empathic stance of Rogers, operationalized it, and added directionality by fully incorporating the skills of the

therapist in initiative dimensions of action. Those who had higher levels of functioning on the Carkhuff Responsive and Initiative Scales (Carkhuff, 1969) had more successful outcomes. This has been validated in counselling, counsellor training, counselling by paraprofessionals, rehabilitating delinquent adolescents, and parental influence. (Webster, 1972) With such a rich theoretical background, soundly based in a concern for people, it was not long before educators began to utilize the Carkhuff approach.

Hefele (1971) trained in interpersonal skills some graduate students who were preparing to be teachers of the deaf. They were evaluated by their supervising teachers and by the achievement of the students that they taught. Significant benefits were claimed for the graduate students in that their interpersonal ability was improved and their students' academic achievement was superior to that of a control group.

Kratochvil (1968) investigated the levels of functioning of teachers and parents as measured by the Carkhuff Scales. The investigation proposed that there would be a relationship between the level of functioning of the adults and the performance of the children on several measures. The results were not significant as there were very few teachers or parents functioning at minimally effective levels or higher.

One of the main exponents of the Carkhuff Model in classroom education has been David Aspy (Aspy, 1972). In one process study he found that the teachers who exhibited a higher positive regard for students elicited higher levels of cognitive functioning from their students. In another study the teachers who were more accurate with empathy, were more genuine, and showed more respect had students who had higher gain scores in academic achievement. Aspy raised the possibility from this naturalistic study that the teachers with the highest skill in interpersonal functioning may also be the most knowledgeable teachers. He also raised the question of the appropriateness of empathy, genuineness, and respect for all age levels as his study was done only with the third grade. Reed (1961) had findings that seemed to show that older pupils did not profit from warmth.

Aspy (Aspy and Hutson, 1972) also studied the results of teachers who utilized students' goals as the basis for classroom activity. Those teachers who were able to do this ranked the highest on the Carkhuff Scales and had more instances in the classroom of student

initiated ideas.

In an experimental study to measure the effect of systematic human relations training of primary teachers upon their classroom performance (Berenson, 1971) it was found that those who experienced an effective interpersonal relationship and who were taught the components of effective interpersonal skills changed their classroom behaviour toward being more effective. Those who were only taught but did not have the experience did learn to discriminate effective from non-effective teaching behaviour. However, they did not change their own teaching behaviour toward being more effective. They remained where they were before they were taught. As measured by the Flanders Interaction Analysis Category System the first group of teachers had a significantly higher indirect to direct ratio, meaning that they were able to expand student participation through accepting pupil feelings, praising and using pupil ideas, and questions more than the second group of teachers. (Flanders, 1970)

In general the findings are that teachers who respond most accurately to the child are able to involve the child to a greater extent in cognitive attainment. The studies cited have provided a good indication of the necessary skills. One of the major questions that remains is, are these skills inborn, unlearned, or accidents of "personality," or are these skills such that with appropriate training, they can be acquired by teachers and used to the same effect as exhibited by Aspy? A similar problem was also raised by Kratochvil (1968) who failed to find enough high level functioning people. He saw the need to train for improved functioning as measured by the Carkhuff Scales. With this increase in skill, there may be an increased differentiation on outcome measures that can be attributed to the subject's level of functioning.

## THE HYPOTHESES

In summary the interrelations that have been developed are the following:

1. the child's parents and significant others train the child physically and cognitively toward the level that they have reached
2. the child's parents and significant others shape the affective field of the child, creating positive and negative valences to behaviours, some of which affect cognitive performance
3. upon entry to formal education, the child's potential for cognitive achievement has been developed to limits set by his environment
4. the teacher in trying to improve the cognitive development of the learner is more effective if he is able to respond accurately to the affective field of the learner and to diagnose the cognitive level of development of the learner.

Proceeding from the information that has been determined as to the relationship of the cognitive and affective nature of the learner, several major questions have been advanced. In summary these questions are:

Can teachers be trained in skills that will allow them to respond accurately to the learner?

Will students of the above teachers demonstrate increased cognitive achievement in like manner to those students of teachers who were naturally able to respond accurately to the learner?

Is there an age level after which learners do not profit from teachers who are able to respond accurately?

### Hypotheses

It is assumed that teachers will be able to gain the skills that will enable them to function in the classroom at higher levels as measured by the Carkhuff Scale for Effective Helping.

The major hypothesis and the subsidiary hypotheses arise from the preceding questions, and obtain their directionality from the studies and results that have been gathered so far.

### Major Hypothesis

The level of accurate responding and initiating behaviour of the teacher will be positively associated with differential student academic achievement.

### Subsidiary Hypotheses

H1 Teachers with high level responding and initiative behaviour will have students who achieve greater gains in academic achievement than students of teachers who do not exhibit high level responding and initiative behaviour.

H2            There will be no difference in the gain in academic achievement for students of all ages who have teachers who have equally high level responding and initiative behaviour.

H3            There will be no difference in the gain in academic achievement for male and female students who have teachers who have equally high level responding and initiative behaviour.

## RESEARCH DESIGN

The ideal experimental design for testing these hypotheses would involve two groups of teachers who are equivalent in all characteristics save that of the ability to attend and respond to children. A random assortment of pupils would be given to each group of teachers. After the pupils have experienced their assigned teachers for a length of time, say one year, the performance of the pupils would be assessed and compared. This would give a direct test of the effect of the research variables. Unfortunately, there are several conditions within our society that prevent this design from being implemented. These conditions are the non-availability of a pool of teachers with high level attending, responding, and initiating skills, the non-random sorting of pupils and teachers into schools and pupils to teachers, and the non-equivalence of teachers on a wide variety of attitudes, skills, and experience. The effect of these conditions upon the research design will be considered.

The skill of accurately responding to the pupil's frame of reference requires the teacher to observe non-verbal messages from the pupil, to listen to what the pupil says, and then to respond empathically by identifying how the pupil feels and the cause of his feeling. The absence of this skill in the classroom has been noteworthy. Flanders designed a classroom observation schedule that was to describe the interactions in a classroom. One category was the acceptance or clarification of feelings. This category has been involved in less than one per cent of all classroom interactions. It appears that the description of the classroom as an "emotional desert" is accurate. (Flanders 1960, 1970)

Kratochvil (1968) designed extensive research to identify those teachers and parents who were able to respond empathically. He then wished to correlate these adults' level of empathic responsiveness with objective measures of their children's performance. However, there was a low, limited range of performance among his adults. There was an absence of a readily available group of adults, teachers and parents, who could respond accurately to their charge's world. Because of this attenuation of skill levels, he was unable to detect significant differences in the performance of the children as a function of those

adult behaviours.

Although there has been little attempt to identify the inter-personal skill level of teachers in New Zealand, it may be expected that there will be a similar lack of empathic ability among teachers. To test the hypotheses it will be necessary initially to train a group of teachers in responding skills. Measures of the teachers' initial base line classroom behaviours in attending and responding skills will be taken, followed by a training input, and a subsequent measure of the execution and maintenance of the relevant skills. This group of trained teachers will have to be compared with a group of untrained teachers to test their initial equivalence or variance in the target behaviours and to ascertain whether the training made a difference. The children of all classes will have to be compared on relevant achievement measures.

Whenever there is an intervention into a productive environment, that is perceived to be helpful, be it at the Hawthorne Western Electric Plant or in a classroom, it appears that there is an increase in productivity. (Kerlinger, 1964) To test whether the hypothesized factors are the important ones it is also necessary to intervene with another group of teachers, preferably in a way that is seen to be educationally helpful but which has been shown not to be a source of significant improvement. These teachers and students will also have to be compared to those of the other two groups.

The design of the research is shown in Table 1.

TABLE 1  
The Research Design

	Pre-measure	Experimental Treatment	Placebo Treatment	Post-Measure
Teachers "A"	Yes	Yes	No	Yes
Teachers "B"	Yes	No	Yes	Yes
Teachers "C"	Yes	No	No	Yes
Students "A"	Yes	No	No	Yes
Students "B"	Yes	No	No	Yes
Students "C"	Yes	No	No	Yes

This design is similar to the Pretest-Posttest Control Group Design of Campbell and Stanley (1963).

The pupils will only experience a pre- and a post-test measure of their performance. There will be no direct intervention by the experimenter with the pupils. Any difference that will be detected will be attributed to the teacher's behaviour.

Another deviation from the ideal experimental design is that students are not randomly assigned to their teacher. There is a sorting by age initially with additional considerations of performance. If a student is doing exceptionally well or poorly, he might be moved up or down a class. An additional source of systematic variation is that schools draw their pupils from the immediate community which is not representative of the entire New Zealand population. There are concentrations of racial types, occupational groups, and educational levels that influence the school's composition.

Studies done to investigate pupils in New Zealand, such as that done by Dawson and Hallinan (1969), have shown that even in two schools, only a short distance apart there have been wide variations on input factors, such as family size and parent's occupation, and output factors, such as pupil reading ability and general intelligence. An initial reading of such a study can leave the impression that it would be impossible to compare students from one school with the students from another school. However, even though the students at one school went in lower and came out even lower again relative to the students of the other school, the standard score difference between the two did not change. The school did not lessen the difference between the two kinds of pupils, but neither did it increase the difference. (Humphreys, 1974) It should be possible to compare pupils even from widely different backgrounds if their initial levels are noted. One method would be to compute a gain score for each child based on two tests of the same material. The gain scores of all the pupils at one school could be compared with the gain scores of the other school's pupils. An alternative would be to analyze the differences through the analysis of covariance using the initial level of performance as the covariant. The latter method is preferred by Stanley (1968) and Kerlinger and Pedhazur (1973).

While there will be enough pupils to level off the impact of an exceptional pupil, how can all of the factors that might influence teacher performance be held constant except for those factors that are to be manipulated? Experience, initial teacher training experience, sex, marital status, health, in-service courses, hobbies, school climate

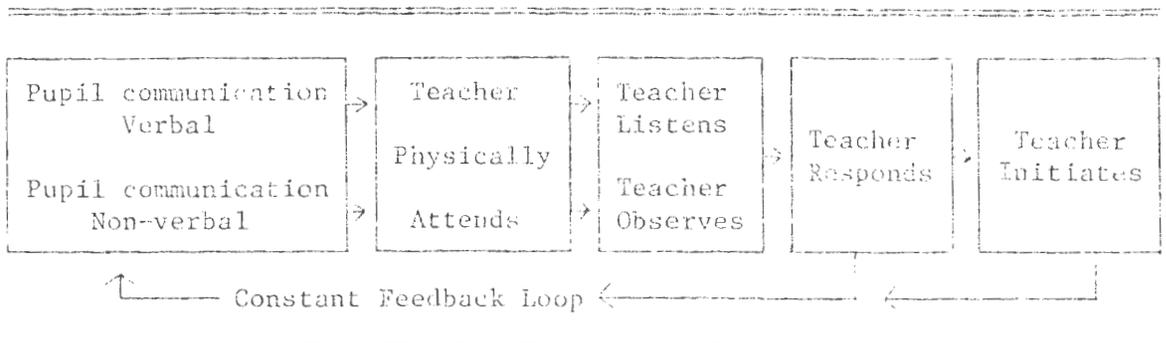
as set by the principal, size of family of orientation, and the list goes on of possible factors that can be hypothesized to be of critical importance. Teachers are not placed in schools randomly. These are illustrations of some of the limitations of an experiment conducted in a naturalistic setting. Instead of being a laboratory setting where more of the factors could be controlled, this will be a field experiment in which as many of the factors as possible will be identified and tested to identify that which does make a difference. The major hypothesis identifies several factors that have strong indications that they are independent variables in teacher delivery of the aims of education.

So far there have been references to certain skills that possibly make a difference; namely accurate empathy, attending, responding to feelings and meanings, and initiating. These will now be defined and operationalized.

## INSTRUMENTATION

To be able to respond in a way that demonstrates that the teacher knows the pupil's frame of reference logically requires a progression of behaviours. During the student's emission of a communication, verbal and/or non-verbal, the teacher must be attending to the communication physically. The teacher must observe the non-verbal behaviour and listen so that he may recall what has been said. Based on this information received the teacher can show that he understands by responding to what has been communicated and subsequently initiating a course of action based on his understanding of the pupil. This can be seen in a flow chart that is Table 2.

TABLE 2  
Pupil-Teacher Communication



### Scale for Attending

An increasing amount of information has been gained on the effect of certain non-verbal behaviours on liking and other approach attitudes such as warmth and respect. Mehrabian (1972) ranked the following behaviours that show friendliness: touching, decreasing distance, leaning forward, maintaining eye contact, and orienting one's body squarely toward the other. However, Haase and Tepper (1972) found eye contact and forward lean to be more highly related to empathy than was distance. The least effective communicator-listener was one who was farthest removed, leaned backwards, had no eye contact, and had low verbal empathy. Even high verbal empathy was unable to negate the deleterious effects of no eye contact and backwards lean.

Gladstein's (1974) review of 115 studies of non-verbal communica-

tion in counselling concluded that non-verbal communication can be reliably classified and that it does have influence on the counselling process. However, there were few unified conclusions as findings were not replicable across different populations of helpers and helpees.

The complex interactions that are involved in non-verbal communication behaviours are beginning to be investigated. Genthren and Moughan (1977) found that extroverts are threatened by a quiet, upright listener, whereas, introverts appreciate a listener who is silent, and even more so, one who leans forward.

Any attempt to establish particular meaning for non-verbal behaviours will of necessity have to specify the exact context and attempt to obtain measurement in naturalistic settings. A forward lean with eye contact can be interpreted as threatening or indicative of increased attentiveness based on the amount of trust that the initiator has earned.

Carkhuff and Pierce (1975) in spelling out the training model concentrate on three major indicators of attention:

Squareness: the orientation of the body and the head so that the left shoulder of one person is directly opposite the right shoulder of the other and his right shoulder is directly opposite the other's left shoulder;

Eye Contact: the maintenance of direct, mutual gazing;

Lean: the inclination of the trunk of the body from the hips, forward toward the other.

The proposal is that a helper is most effective by beginning in a square position, with steady eye contact, and with a slight forward lean. With the establishment of trust as evidenced by high level empathy, the helper can increase his lean toward the other.

As most of the non-verbal communication studies have been based on counselling settings between adults (Gladstein, 1974) very little has been reported on non-verbal communication between adult and child. The scale shown in Table 3, is based on Carkhuff's approach to training and incorporates an additional factor of level which means that the teacher and the pupil have their heads on the same level. The

lowest numbers refer to those physical behaviours which show the least amount of attention being given, while the higher numbers refer to more intense attention being given by the teacher.

TABLE 3  
Scale for Teacher Attending

Level	Behaviour
1.0	The teacher is not squarely facing the pupil, and has no eye contact.
1.5	The teacher is squarely facing the pupil and has no eye contact.
2.0	The teacher is squarely facing the pupil and has intermittent eye contact.
2.5	The teacher is not squarely facing the pupil and has steady eye contact.
3.0	The teacher is squarely facing the pupil and has steady eye contact.
4.0	The teacher is squarely facing the pupil and has steady eye contact and leans toward the pupil.
5.0	The teacher is squarely facing the pupil, has steady eye contact, leans fully toward the pupil, and is on the same level as the pupil.

A rating is given when the teacher is with an individual no matter for what length of time. It is necessary for both the teacher and the pupil to be visible to the rater. If the teacher's and the student's posture is stable during the interchange, one rating is given. If their posture changes, an additional rating is given. For example: the teacher approaches and stands next to a student who is also standing, working on a project. Both of them look at the project. This would be rated a "1.0". While they are talking, the teacher turns and faces the pupil who also faces the teacher. Both of them maintain a steady eye contact. This would be rated a "3.0".

"Facing squarely" means that the teacher and the pupil are positioned so that the right shoulder of one is directly opposite the left shoulder of the other.

"Eye contact" is present when the teacher looks at the pupil's

eyes and the pupil looks at the teacher's eyes.

"Lean is present and relevant when the teacher and the pupil face each other squarely, maintain steady eye contact, and the teacher inclines his body from the waist toward the pupil.

"Being on the same level" means that the head of the teacher and of the pupil are level.

#### Scale for Teacher Responding and Initiating

While there are specific skills involved in being able to listen and observe well, these are not visible to outside observers. Therefore, these skills can be taught but not observed directly. Their validation can be achieved only through verbal responses at the time or through later recall. The skills of responding are more amenable to observation.

Many psychologists and psychiatrists consider their profession to involve an educative process whereby their clients are helped by learning how to overcome some problem or how to live in a more satisfying manner. (Eysenck, 1960; Ellis, 1962; Yablonsky, 1965; Krumboltz, 1966; Carkhuff and Berenson, 1967) One could consider the more successful practitioner to be the one who was the best teacher. Several theoretical approaches have gained large numbers of followers who believe that their way is the best way. This has led to frequent arguments between the different camps with the result that the primary goal of being effective helpers, not measured by how well a theoretical model has been followed but in terms set by the client, has often been forgotten. This has led to very potent criticisms of the psychotherapeutic field. (Eysenck, 1965; Levitt, 1963; Truax and Carkhuff, 1967; Carkhuff and Berenson, 1967)

Charles Truax and Robert Carkhuff (1967) resolved to take an atheoretical approach to the problem of counsellor effectiveness. They investigated the behaviours of the therapist that were associated with gains for the client. Some actions were found to be more effective than others. There were effective helpers in all theoretical groups. After determining the behaviours for effective helping, Carkhuff (1969) developed scales that were an attempt to operationalize the definition of levels offered by helpers.

The first scales measured the helper's empathy, respect, genuineness, confrontation, and immediacy and self-disclosure. The higher the

level on each of the scales, the more helpful the helper would be. Subsequent research showed high intercorrelations among the scales and pointed out areas of possible integration. The two main areas that were shown to be critical were the ability of the therapist to respond accurately to the experience of the client and the ability of the therapist to initiate from his experience of the client and his own life in a way that would be acceptable and helpful to the client. (Carkhuff, 1972)

Educators soon began to use these scales in an attempt to measure the classroom effectiveness of teachers and as a guideline for the training of effective teachers. (Aspy, 1972; Berenson, 1971; Hefele, 1971; Carkhuff and Pierce, 1976) Hefele's (1971) scales were not operationally defined and depended upon some inference by the rater.

The facilitative teacher is a person who is living effectively and who discloses himself in a genuine and constructive fashion in response to the pupil. He communicates an accurate empathic understanding and a respect for all of their feelings and learning experiences. He communicates confidence in what he is doing, is spontaneous and intense, and is open and flexible in his relationship with them and overtly committed to the welfare of all of the pupils. His facilitative involvement is partly evidenced by the amount of appropriate verbal and physical activity in which he engages with the pupils throughout the lesson.

Rate each segment using the continuum below.

- 1.0 None of these conditions are communicated to any noticeable degree by this teacher.
- 2.0 Some of the conditions are communicated and some are not.
- 3.0 All conditions are communicated at a minimally facilitative level.
- 4.0 All conditions are communicated and some are communicated fully.
- 5.0 All conditions are communicated fully, simultaneously, and continually.

Even with the shortcomings of this particular scale, raters were able to discriminate effective from non-effective helpers.

The latest scale and accompanying trainer's guide (Carkhuff and Pierce, 1975) has incorporated further refinements which will be beneficial in rating helper performance and in the training of helpers because of its focus upon the necessary behaviours included at each level. What is presented in Table 4 is an adaptation of the Carkhuff Scale for Effective Helping that can be used in classroom observation.

TABLE 4

An Adaptation of the Carkhuff Scale for Effective Helping  
for use in Classroom Observation

Based on Carkhuff, Robert and Pierce, Richard (1975)  
Trainer's Guide: The Art of Helping Palmerston North:  
Human Resource Development Press.

Level	Teacher Behaviour
1.0	The teacher responds by ignoring or criticizing the student. The teacher may continue on from his own frame of reference.
1.5	The teacher responds to the general content of the student. The teacher may ask the student to further explain after the student has already given a reply.
2.0	The teacher responds accurately to the specific content of the student with no reference to the feeling of the student.
2.5	The teacher responds accurately to the feeling of the student, but the content is either not present or is inaccurate.
3.0	The teacher responds accurately to the feeling and the content of the student.
3.5	The teacher responds accurately to the feeling and personalized meaning of the student.
4.0	The teacher responds accurately to the feeling of the student and states what the student lacks.
4.5	The teacher responds accurately to the feeling of the student and establishes a personalized goal for the student.
5.0	The teacher responds accurately to the feeling of the student, establishes a personalized goal, and gives a first step of action to reach the goal.

The observer must be able to hear clearly the teacher's and each individual's voices. Seeing the participants will help the rater to add non-verbal clues to the verbal components. While the non-verbal behaviour of the teacher is pre-potent over his verbal behaviour, (Haase and Tepper, 1972) it is hoped that teacher verbal and non-verbal messages are congruent. The scale does not specifically delineate any non-verbal criteria, although Carkhuff in his training model trains body posture and other attending behaviours that are helpful before he trains for verbal responses.

The unit of response that is evaluated is the response of a teacher to a pupil utterance. Each response will be rated and recorded in only one category. To aid in determining the accuracy of the teacher's response the rater must listen to the pupil's reaction for the acceptance or rejection of the teacher's response. A typical flow of classroom talk would be the following:

- (1) teacher talk
- (2) pupil comment
- (3) teacher response
- (4) pupil reaction

The unit that is rated is (3) teacher response. The rater must be able to hear (2) pupil comment and (4) pupil reaction both of which will assist the rater in his judgement of the accuracy of the teacher's response. If the flow continues with (5) teacher response to (4), the rater will determine if there is any additional communication from the pupil, e.g. (6) pupil reaction, and make a second rating of the teacher's responding ability. A change in the pupil doing the reacting does not change what is recorded.

- (1) teacher talk
- (2) pupil A comment
- (3) teacher response
- (4) pupil B question
- (5) teacher response
- (6) pupil B reaction

The rater will code the teacher on each response, (3) and (5). Since there is no reaction recorded from Pupil "A" to (3) teacher response, the rater must infer from the information given by (2), pupil "A"

comment, the accuracy of (3) teacher response.

From level 1.0 to 3.0 the focus of attention is upon whether the teacher has accurately identified the feeling of the student and the reason for the feeling as expressed by the pupil.

Response Level	Feeling of Pupil	Meaning of Pupil
1.0	absent from response	absent from response
1.5	absent from response	partially present
2.0	absent from response	accurately included
2.5	accurately included	partially present or absent from response
3.0	accurately included	accurately included

The levels above 3.0 are attained only after the teacher has expressed to the pupil a level 3.0 response. The higher levels are based upon the addition to the response of important pieces of information that will enable the pupil to understand his situation and act purposefully.

At level 3.5 the teacher is able to respond to what the pupil is saying about himself. A format that is used in training to achieve a 3.5 level of response is, "You feel \_\_\_\_\_ because you \_\_\_\_\_." To determine the accuracy of the teacher's response at all levels there must be an acknowledgement from the student of the correctness of the response.

At level 4.0 the teacher is able to state a deficit in the student's behaviour. The format for training this response takes the form, "You feel \_\_\_\_\_ because you can't \_\_\_\_\_." The deficit may be a lack of information or a larger behavioural deficit that hinders the student.

At level 4.5 the teacher is able to state the student's goal. The goal is the reverse of the problem or deficit first identified in the level 4.0 response. If the deficit is the inability to tell others how he thinks and feels, the goal is for the pupil to be able to communicate clearly with others. The teacher must use his training as an educator and his personal experience to be able to communicate at these additive levels. At this level the teacher sets the goal based upon level 3.0 and 3.5 responses that shows he knows the pupil's world and a level 4.0 response that is acknowledged by the pupil to be the problem that needs correction so that he can continue to grow in skills and develop a more satisfying life.

At level 5.0 the teacher gives the first necessary step of action that the pupil can do to achieve his goal. The teacher analyzes the goal into the intermediate steps necessary and further into the immediate action steps for the student to accomplish his goal. In entering and completing this process the teacher ensures that the student will experience success in identifying his present situation, in identifying his problem area and his goal, and in initiating action to solve his problem. In the future he will be better equipped to be a better problem solver.

This scale gives a static overall rating of teacher responses obtained by adding all of a teacher's response levels together and dividing by the total frequency. Also included is a preferred sequence of responses for the teacher. In the counselling situation effective helpers begin first with approximately six level 3.0 responses. Then they move to 3.5 responses. Depending on the level of the helpee, the counsellor can move to the definition of personal deficits and personal goals. Through the entire process the effective helper constantly goes back to level 3.0 to check that he is with the helpee. A series of very effective responses could be like this:

3.0; 2.5; 3.0; 3.0; 3.5; 3.0; 3.5; 3.5; 3.0; 4.0;  
3.5; 3.0; 4.0; 3.5; 4.5; 3.0; 4.5; 3.0; 5.0; 3.5.

The reality of teaching usually removes the possibility of a teacher talking with one student publically through this many responses. However, the usual level of teaching has been very low (Kratovich, 1968) so that even one 3.0 response would be a great improvement toward recognition of the affective and cognitive interaction that is involved in learning. Also a teacher could move through the sequence more rapidly because he has dealt with the pupil before. A sequence for a teacher and a pupil could go 3.0; 4.0; 4.5; 3.5; 5.0.

There are several things which the scale for effective helping is not. It is not a continuum measuring at level 1.0 classroom activities that are solely toward the achievement of teacher goals to level 5.0 activities that are solely toward the achievement of student goals. Goals are very important and the definition of goals at level 4.5 involves both the teacher and the student in the formation of mutually acceptable goals. Very many teachers have been influenced by a form of humanitarian education that has removed teachers from being responsible.

The Carkhuff approach gets the teacher back into responsibly acting for the benefit of pupils in his classroom as well as respecting the great worth of each individual.

The scale does not place classroom activities on an authoritarian-democratic continuum. In both of these conditions less than optimum performance will be achieved because the different contributions and responsibilities of both the teacher and the pupil are not recognized.

Many educators like to consider classroom activities on a lecture method versus discovery method continuum. Again this is not the dimension uncovered by utilizing this scale. The effectiveness of a teacher can be enhanced under either extreme method if he is able to respond to his students. The discovery method would give more opportunities for student initiative, but without the teacher skills of being able to respond accurately and initiate appropriately, many educational opportunities will be lost. Even if a teacher uses the lecture method, he will be more successful in reaching his goals if he is able to raise his own level of responding to student communications.

In fact it is doubtful that the scale is a continuum of any dimension except that which it purports, namely effective helping. At the low end of the scale are those behaviours which have been found to be least effective. The other end of the scale has those behaviours which have been found to be most helpful. The formation of the scale evolved from the study of counsellor behaviours which were associated with positive outcomes. It is not surprising that there have been outcome studies with the scale predicting the criterion. In teaching settings earlier versions of the scale have been effective in predicting success of student achievement, positive student attitudes, and supervisor ratings of teacher behaviours. (Aspy, 1972; Berenson, 1971; Hefele, 1971) It is expected that this version will also predict with the same power plus having the additional benefit of describing teacher behaviours and processes that will lead to increased learning and a more positive attitude toward learning for the pupils affected.

#### Measures of Pupil Achievement

If the teacher can attend to, respond to, and initiate with the child, there may be improved performance by the pupil in meeting the aims of the school. These aims can be thought of as increased intellectual and interpersonal skills as outlined by the official

syllabuses.

One of the methods of measuring intellectual performance is through standardized testing of general intelligence. Another method is to use a standardized test of curricular achievement. Additionally, there could be limited testing of short, prescribed material. It was decided to use the standardized test of curricular achievement as results from this kind of test would be suitable for broad generalizability directed specifically to the school setting.

The Progressive Achievement Test series is currently being used in most New Zealand primary and intermediate classrooms to attain an indication of the gross level of achievement in the areas of mathematics, reading comprehension, reading vocabulary, and listening skills. The tests were constructed by the New Zealand Council for Educational Research based on the official syllabuses and "should provide a reliable and valid estimate of a child's current mathematical achievement."\* (Reid and Hughes, 1974)

While it is unwise to administer such tests to evaluate the progress of an individual pupil after only a short interval because the reliability of gain scores for individuals is only  $r=.20$ , "it is usually satisfactory for comparisons of group performance over a period of time."\* (Reid and Hughes, 1974) As the tests will be evaluated in terms of group performance and will be administered at the beginning and the end of the school year, it is hoped that these tests will give the best possible indicators of pupil performance toward achieving some of the aims of the school.

#### Measure of Pupil Attitude

How does the child view the source of much of his learning? To obtain reliable and valid indications of a child's intellectual performance in school is only part of the measurement of outcomes. The attitude of the child toward the teacher can be obtained through lengthy questionnaires, private interviews, systematic differential checklists, unobtrusive measures, or sentence completions. All methods of obtaining a measure of the attitude of the pupil towards the teacher can be confounded and invalidated through pupil fear to answer in a full and truthful manner or through pupil compliance to the wishes of others. Even with these shortcomings, the obtaining of some indication of pupil attitude towards the teacher is considered

\* Similar statements are present in the manuals for vocabulary, comprehension, and listening tests.

important enough to attempt to obtain. It must be able to be administered easily to young children and be easily and clearly interpretable. It was decided to use a very simple sentence completion form, "I think my teacher \_\_\_\_\_." This will be embedded amongst other completion frames. To help remove the inhibition of fear the response will remain anonymous. The responses will be divided into three categories: those which are positive toward the teacher, those which are neutral toward the teacher, and those which are negative toward the teacher. A similar sentence stem and procedure was used by Webster (1972) who found that the attitudes of youths was highly correlated with objective measures of adjustment and achievement.

## TRAINING PROCEDURES

### Experimental School

The training sessions with the experimental school teachers aimed at producing teachers who could pay attention with their pupils at an average of 2.5 on the Attending Scale and who could respond to their pupils at an average of 2.5 on the responding scale.

The ten sessions had the following aims and activities:

- Session 1:     Goal: Teachers who can pay attention to each other while standing and sitting at level 3, 4, and 5.  
 Methods: Point out the contributions of eye contact, being square, leaning toward the speaker, and keeping the head erect in paying attention. Practice with each other the level three behaviour while standing and sitting. Then, increase body lean toward each other. Respond to their experience and maintain attention.
- Session 2:     Goal: Teachers who are able to listen accurately and are able to repeat verbatim student talk which lasts for ten seconds.  
 Methods: Discuss distractions that hinder teachers from hearing what students are saying, both from external and internal causes. Arrange the environment to help reduce distractions so that the teacher can hear. Suspend judgements of good and bad. Give examples of student talk which the teachers attend, listen, and then repeat verbatim.
- Session 3:     Goals: Teachers who are attending and listening in their classrooms and teachers who are able to observe student behaviour.  
 Methods: Attend and respond to each other while recounting effects in the classroom of paying attention and listening. Point out methods of observing; look at the observable, do not infer very far, do not hurt, look for behaviours that give an indication of pupil's energy level and

feelings, also note any behaviours that do not match. Observe a role play for thirty seconds and report observations.

- Session 4: Goal: Teachers who are able to identify feelings of pupils.  
 Methods: The teachers role play a student talking that they have experienced this week. The others attend, listen, observe, and then write down what was said and then identify the feelings of the pupil.  
 Discuss the major families of feelings: happy, sad, angry, fear, and confusion along with words that would give different levels of intensity of feeling.  
 After another role play, respond by identifying the feeling by using the framework, "You feel \_\_\_\_\_."
- Session 5: Goal: Teachers who are able to respond to feelings and their cause.  
 Methods: Respond to their experiences this past week of identifying and responding to pupil's feelings.  
 Discuss the causation of feelings.  
 Role play teacher doing and saying a pupil behaviour. The other teachers attend, listen, observe, and respond by identifying the feeling and its cause. A helpful framework is, "You feel \_\_\_\_\_ because \_\_\_\_\_."
- Session 6: Goal: Teachers who are able to respond to feelings and their cause.  
 Methods: Pair up one teacher with another. One teacher tell of some incident in the past week that was important for him. The other teacher respond only with how the teacher feels and what causes this feeling. If possible, work toward three extended interchanges.  
 Give pupil talk examples which the teachers respond to.
- Session 7: Goal: Teachers who are able to justify the inclusion of feelings in the classroom.  
 Methods: Present a model of teaching: teacher goals to preparation of the environment to observing, listening, responding, identifying deficits to teaching to observing, listening, responding, identifying deficits, establishing goals to teaching.

- Session 8: Goal: Teachers who can evaluate their own teaching behaviour.  
Methods: Teachers observe their teaching delivery as recorded on video-tape and use the Attending Scale as an aid to evaluate.
- Session 9: Goal: Teachers who attend and respond to pupils.  
Method: Teachers teach a five-ten minute lesson to three-five of their students. The other teachers observe the success or not of attending and responding.
- Session 10: Goal: Teachers who continue to attend and respond to their pupils.  
Method: Present the overall Effective Helping Scale for teachers.

#### Placebo School

The treatment sessions of the placebo school had as their aim increased variability of question types asked by teachers in the classroom. Also it was desired to have teachers to prepare objectives for their students utilizing behavioural performances based on the different levels of complexity as identified by Bloom. The additional aims were for the presentation to be seen to be of educational value but not to improve or hurt the pupils' performance. As summarized by Dunkin and Biddle (1974) the taxonomy developed by Bloom and others for cognitive objectives would meet the requirements for this part of the experiment.

In order to have a basis for constructing clear and unambiguous goals for teachers to use Bloom (Bloom et al, 1956) developed The Taxonomy of Educational Objectives. The taxonomy was sub-divided into the cognitive domain, the affective domain, and the physical domain. Only the cognitive domain will be considered here. The writers assumed that cognitive functioning could be classified along a single continuum of complexity. At one end of the scale would be very simple cognitive functioning. At the other end would be activity that would exhibit greater complexity. Another assumption was that the scale was a Guttman Scale. The exhibition of higher, more complex functions implied mastery of the more simple functions. Several other hidden assumptions could be teased out from their writing, but one very

prominent set of assumptions was present. Teachers do not teach their pupils to think. If teachers would be aware of the more advanced levels of cognitive activity and actively incorporate this awareness into their classroom behaviour, their students would learn to use more complex cognitive behaviour, hence they would be able to think. Bloom and the very large contingent of educators who worked with him were clearly trying to improve pupil learning.

Their first intentions were that teachers would use the taxonomy to construct educational aims and that they would then develop test items based directly upon their aims. It was assumed that the classroom instructional behaviour of the teacher would be heavily influenced by their personal aims. To see how successful they were the continuum will be presented, then several adaptations of the taxonomy for use in classroom observation will be examined, research findings will be summarized, and the implications for use and future development will be extrapolated.

The Bloom taxonomy has six major categories each of which are sub-divided.

1. Knowledge - This involves the recall of specific bits of information, the recall of methods, or the recall of principles or theories.
2. Comprehension - This is the lowest level of understanding. With this ability the person is able to translate a communication from one language or form to another. He can interpret an entire work or extrapolate from the information given.
3. Application - The person can demonstrate the use of abstractions in concrete situations.
4. Analysis - The breakdown of a communication into its elements that leads to clarification of the communication is the central skill. In doing this the elements, the relationship

of the elements, or the organizational principles of the elements are made clear.

5. Synthesis - This is the putting together of elements to form a whole. This skill could be demonstrated by a unique communication, a plan of operations, or the derivation of a set of abstract relations.
6. Evaluation - This involves making a judgement about the value of material or methods to meet a given purpose whether measured by internal or external criteria.

This taxonomy was first published in 1956 and gradually gained an audience of users. However, it was not until 1968 that a category system for classroom observation based on the Bloom taxonomy was published. In fact 1968 saw two systems published.

The Teacher-Pupil Question Inventory (Davis and Tinsley, 1968) is a category system that expects a coding by the observer for each question asked by the teacher or the pupil. In doing this the instrument focuses on only one kind of incident in the flow of classroom events. The observer must be able to hear both the pupil's and the teacher's voices. Whether this is done through audio or video recording or by a live observer is acceptable as long as the rater can hear both voices. Every incidence of a question would be put into only one category with two pieces of information: who asked, teacher or pupil, and what was the level of complexity as determined by the scale. In addition to the measurement of the level of cognitive difficulty the TPQI includes two additional categories that will give a more complete picture of classroom activity. The first is the category of Affectivity and the second attempts to cover those areas broadly defined as questions of Procedure.

TABLE 5  
The Teacher-Pupil Question Inventory

---

---

1. Memory - The questioner expects the recall or recognition of information.
  2. Translation - The questioner expects a change in the form of the information. (Symbolic or Linguistic)
  3. Interpretation - The questioner expects to be shown the relationships between various pieces of information.
  4. Application - The answerer is expected to solve a realistic problem requiring the identification of the relevant parts and the use of appropriate skills and knowledge.
  5. Analysis - The questioner expects explicit attention to the relationship between the ideas presented.
  6. Synthesis - The questioner expects an original, speculative, or creative answer.
  7. Evaluation - The answerer is expected to give an effective judgement based on explicit criteria.
  8. Affectivity - The questioner expects a statement of feeling, emotion, or opinion without a standard of appraisal.
  9. Procedure - The questioner expects a statement dealing with classroom organization, instructional management, or student behaviour.
-

As can be seen from the TPQI, there have been additional changes from Bloom's original taxonomy. Bloom's level two, Comprehension, has been expanded to include Translation at level two and Interpretation at level three on the TPQI. This division was included in Bloom's original work under a single category.

The categories as developed by Bloom and utilized by classroom observers have an attractive theoretical appeal. Certainly thinking can be simple or it can be more complex. A more advanced mode of thought would be more complex. Teachers would be well advised to increase the complexity of their questions and the answers that they expect from their pupils. By doing this many benefits will follow. This is sound thinking if all of the assumptions of the taxonomy are true. However, the research attempted so far has drawn attention to weaknesses in the original conceptualization.

The first problem to be confronted in using the taxonomy is the lack of discreteness of the categories. If a teacher asks his pupils to draw a picture of improvements to the human body, is this translation, application, or synthesis? If a standard one pupil is asked to add two plus one, for him it could be a complex synthesis of the concepts that have been developing, whereas, for another standard one pupil it could be a simple statement of memory.

To the question, "What do you think about slavery?" a pupil's answer, "I don't like it," could be categorized as a statement of affectivity if a standard of appraisal is lacking or evaluation if the student recognizes the humanity of the slave and presupposes the universality of her feelings and values. The focus of the TPQI, is upon the question as the title rightly leads one to suspect. Does such a question expect an answer of evaluation, affectivity, analysis, or whatever level of complexity? The categories in the TPQI and in Bloom's taxonomy are not discrete entities. Any attempt to categorize classroom behaviour using these scales will involve a high amount of inference. Frequently, the coding will be based upon the answer's level of complexity. This negates the intention of the questioner.

Another problem has been with the assumption of the unidimensionality of cognitive functioning. Thought can be conceived to differ on the dimension of complexity. Another variant is abstractness. Simple thought can be very concrete or very abstract. Likewise, complex thought can be quite concrete or abstract. Solomon and Wood (Solomon

and Wood, 1970; Solomon, 1970) reported that an increase in the complexity of teacher's questions was accompanied by a decrease in abstractness and an increase in concreteness. This indicated that teachers could increase the complexity of their questions, but to meet their pupils, they lowered the level of abstractness to a more concrete level.

Clearly, one could expect to encounter difficulties when trying to utilize the TPQI in classroom observation. Most research using this instrument reported acceptable levels of reliability. One (Davis and Tinsley, 1968) reported "almost unanimous" agreement. While the training of raters involved in the mechanical recording of categories has generated satisfactory levels of reliability, one cannot help question the validity of a process of observation based upon the Bloom taxonomy.

One very similar category system is the Florida Taxonomy of Cognitive Behaviour (Brown et al, 1968). As can be seen in Table 6, its fifty-five categories plus whether the teacher or the pupil is the emitter, makes this a very complex instrument. This complexity necessitates a recording system that can be stopped after each codable bit to allow adequate time for the trained rater to code. Live classroom observation would be prohibited. A further variation is that the FTCB is a sign system. In a six minute block of time any incidence of a category is simply ticked off once whether it has occurred once or more frequently. While the number of categories has increased and extended behavioural information has been given to aid in categorization, similar criticisms as raised against the TPQI are still relevant.

TABLE 6

The Florida Taxonomy of Cognitive Behaviour

- 
- 1.10      Knowledge of Specifics
1.      Reads
  2.      Spells
  3.      Identifies something by name
  4.      Defines meaning of term
  5.      Gives a specific fact
  6.      Tells about an event
- 1.20      Knowledge of Ways and Means of Dealing with Specifics
7.      Recognizes symbol
  8.      Cites rule
  9.      Gives chronological sequence
  10.     Gives steps of process, describes method
  11.     Cites trend
  12.     Names classification system or standard
  13.     Names what fits given system or standard
- 1.30      Knowledge of Universals and Abstractions
14.     States generalized concept or idea
  15.     States principle, law, theory
  16.     Tells about organization or structure
  17.     Recalls name of principle, law, theory
- 2.00      Translation
18.     Restates in own words or briefer terms
  19.     Gives concrete example of an abstract idea
  20.     Verbalizes from a graphic representation
  21.     Translates verbalization into graphic form
  22.     Translates figure statements to literature statements
  23.     Translates foreign language to English or vice versa
- 3.00      Interpretation
24.     Gives reason
  25.     Shows similarities, differences
  26.     Summarizes or concludes from objects of evidence
  27.     Shows cause and effect relationships
  28.     Gives analogy, simile, metaphor
  29.     Performs a directed task or process

The Florida Taxonomy of Cognitive Behaviour continued

4.00 Application

- 30. Applies previous learning to new situations
- 31. Applies principle to new situation
- 32. Applies abstract knowledge in a practical situation
- 33. Identifies, selects, and carries out processes

5.00 Analysis

- 34. Distinguishes fact from opinion
- 35. Distinguishes fact from hypothesis
- 36. Distinguishes conclusions from supporting statements
- 37. Points out unstated assumption
- 38. Shows interaction or relationship elements
- 39. Points out particulars to justify conclusions
- 40. Checks hypotheses with given information
- 41. Distinguishes relevant from irrelevant statements
- 42. Detects error in thinking
- 43. Infers purposes, points of view, thoughts, feelings
- 44. Recognizes bias or propaganda

6.00 Synthesis (Creativity)

- 45. Reorganizes ideas, materials, process
- 46. Produces unique communication, divergent idea
- 47. Produces a plan, proposed set of opportunities
- 48. Designs an apparatus
- 49. Designs a structure
- 50. Devises scheme for classifying information
- 51. Formulates hypothesis, intelligent guess
- 52. Makes deductions from abstract symbols, propositions
- 53. Draws inductive generalization from specifications

7.00 Evaluation

- 54. Evaluates something from evidence
  - 55. Evaluates something from criteria
-

Even with increased categories and extended information about the categories, the problem of overlapping categories remains as it did with the TPQI. To try to develop discrete categories that are usable by teachers the Educational Testing Service (no date) drastically condensed the number of categories to only three with only minimal definition of the categories.

Remembering which includes recall of facts, rules, procedures, routine manipulation, and reproduction.

Understanding which includes classification, application, and translation.

Thinking which includes analysis, generalization, or evaluation.

The main purpose of the Educational Testing Service in developing this approach has been to meet Bloom's original intentions, to aid in the formation of educational objectives and in the construction of evaluation instruments. It has not been applied to classroom observation, although it is expected that there would be difficulties encountered similar to those with the more extensive systems.

While there remains a dissatisfaction with the scale, several research findings utilizing the scale are interesting. Teachers and pupils can be trained to use more high level questions. However, a direct increase in the level of teacher discourse will not have the desired effect of raising pupil level of performance. Indeed, with disadvantaged youths higher levels of teacher behaviour were negatively associated with pupil gain. (Ragosta et al, 1971) Additionally, pupils use higher levels of cognitive behaviour the more the teacher relinquishes his authority. (Wood, 1970) Whether there are positive gains by pupils after being exposed to a teacher who has higher levels of cognitive functioning is very much in doubt. In fact students of high level cognitive functioning teachers did less well on analysis type questions than students of lower cognitive functioning teachers. (Rogers and Davis, 1970)

There is evidence that teachers are asking questions and teaching using various levels of cognitive difficulty. There is also evidence that teachers can increase the complexity of their teaching behaviour as measured by one of the Bloom based scales through training. There is evidence that the behaviour of pupils closely

resembles that of the teacher. However, criteria of performance on achievement tests and other objective measures have not supported any claim that raising the level of cognitive discourse in the classroom will improve the ability of pupils to think. This is indeed a disappointment to be dealt with. Possibly the criterion measures do not weigh the gains effected by the intervention. Possibly a combined study of complexity and abstractness would turn up desired teaching behaviours. Possibly the Bloom hierarchy is not true and possibly the hierarchy developed by Gagné (1970) reflects reality more accurately. The information and questions raised by developmental psychology must play an important part. Learning sequences that are effective must also play a part in the complex universe of the classroom. One approach that is hierarchical and sequential is that of Facts, Concepts, Principles, Values, and Programmes. (Carkhuff, 1975 and 1976) Indeed the question originally raised by Bloom and by many educators still stands. What can be done to improve the ability of pupils to think?

The actual placebo treatment was designed as follows:

#### Placebo School

- Session 1:           Goal: Teachers who are aware of levels of thinking.  
                       Methods: Discuss, how do you know whether somebody can think? The ability to solve problems, more than parroting.  
                       Introduce hierarchy of Remembering, Understanding, and Thinking.  
                       Distribute the paper on Bloom Taxonomy that defines the six levels.
- Session 2:           Goal: Teachers who are able to ask questions at the first three levels of the Bloom Taxonomy.  
                       Methods: Define and give examples of Knowledge, Comprehension, and Application. Teachers develop questions from their own units that they are teaching illustrating each level. The other teachers identify the level.

- Session 3: Goal: Teachers who are able to ask questions at the last three levels of the Bloom Taxonomy.  
Methods: Respond to their experiences of asking questions this past week.  
Define and give examples of Analysis, Synthesis, and Evaluation.  
Teachers develop questions from their own units that they are teaching which illustrate each level. The other teachers identify the level.
- Session 4: Goal: Teachers who are able to develop questions spontaneously for all six levels.  
Methods: Discuss students who have difficulty understanding something. To help solve their deficit identify what level you want them to be at, go to lower levels to ascertain the pupil's lack. Give examples.
- Session 5: Goal: Teachers who can turn the categories into activities.  
Method: What is a person doing when he is operating at each level?  
Develop an active verb list for each category that demonstrate that the pupil can work at that level.
- Session 6: Goal: Teachers who can prepare behavioural objectives for each level of cognitive complexity.  
Method: Prepare behavioural objectives for each level of the taxonomy that gives the desired terminal performance, the conditions under which it is performed, and the standard of acceptability.
- Sessions 7 and 8: Goal: Teachers who are able to identify their own level of questions.  
Method: Teachers observe their own teaching performance from video recording and identify the level of complexity of each question.
- Sessions 9 and 10: Goal: Teachers who are able to teach a small group of their children using the various levels of questions.  
Method: Teachers teach 3-5 children from their own class for ten minutes. The other teachers observe and identify the question levels and note the effects in the students to questions of different levels.

## PROCEDURE

Once the major programmes were prepared, three schools were needed which were roughly comparable, teaching under one administration new entrants to Form II with at least six teachers from Standard Two to Form II. A Rural School Advisory Officer from the Wellington Education Board suggested schools that could meet these requirements.

The three schools were assigned to treatment conditions at random. It was assumed that the students and teachers at these three schools were drawn from the normal population of New Zealand students and teachers and that as such their variances were assumed to be equal. Initially, each teacher and each pupil had equal probabilities of being assigned to any treatment. However, as the assignment of treatment conditions was to the group of teachers in one school, there was no possibility of randomly assigning individual teachers or pupils to treatments as the teachers and pupils in one school could not react to the treatment independently of the other teachers and pupils in the same school who would be under an alternative treatment.

For the four assumptions which are necessary for using the fixed effects model for the analysis of variance (Glass and Stanley, 1970) the present field experiment has met the first three assumptions of randomization from a normal population with equal variances for the three schools, the fifteen teachers, and the almost 400 pupils. The difficulty arises with the fourth assumption that the samples are independent. At the level of schools there was a measure of independence as the schools knew that there were other schools involved, but as far as the experimenter is aware they had no communication with each other related to the treatment conditions. The teachers within one school were aware of the treatment that they were under, in fact they received training together and practiced the target behaviour together. Therefore, one could say that the teachers were not independent replications. However, the treatment effects did not have a unidimensional effect on the teachers in one school. It will be shown below that some teachers changed while others did not. Possibly, it can be assumed that although the teachers interacted within one school, they responded in their classroom independently.

The decision as to whether or not the students are independent units or dependent is more difficult to ascertain. The students within one classroom have effects upon each other. Glass and Stanley (1970) give an example of an obnoxious, distracting member of a class who can effectively disrupt and impair the learnings of the entire class. Because the inclusion of such an individual can depress the gains of the entire class, the researcher would be wrong in assuming that the differences between classes was because of teaching methods. These two authors agree with Lingquist (1940 and 1953; see also Lumsdaine, 1962) that any test of difference between intact groups be done on the basis of the means of the groups and the degrees of freedom be based on the number of groups and not on the number of individuals. Thus, the experimental units are the smallest divisions of the experiment that have been randomly assigned to the different conditions in the experiment who have responded independently of each other.

The reality of the classroom and the school contaminate many efforts to achieve the statistical purity that underlie the assumptions for conducting F tests. Students in one classroom are not independent of each other. Surely, they touch elbows and respond to each other. Teachers in one school are not independent of each other as one can witness during breaks and also in staff meetings. One can even suggest that schools being in one educational district interact through teachers attending inservice courses, inspectors, newsletters, journals, syllabuses, sports, and teacher and pupil transfers. Since even the schools cannot be considered to be independent, shall the experimental unit be pushed back to the experimenter with no degrees of freedom? Or possibly the experimenter is part of a larger design in which he is operating independently?

The restriction of the experimental unit responding independently during the duration of the experiment is one which is difficult if not impossible to achieve in a naturalistic setting. One can ignore the pupil interaction and utilize pupil performance as the output unit of statistical analysis. This is done with the danger of increasing the probability of Type I and Type II errors because of individual students who have more influence over the classroom learning environment, either positive or negative, than the teacher. This is a highly suspect argument, especially as the focus in this experiment is the ability of the teacher to respond to and to initiate effectively even with pupils who are disruptive. One can utilize the teacher as the

level of inquiry with the class mean performance as the analytical unit. This choice decreases the probability of detecting changes that are significant because of the conservative degrees of freedom. This choice is attractive and intuitively sensible in that it is teachers that are being compared.

The principals accepted the invitation to participate in the research. During the second visit in the closing week of the 1976 school year, the requirements for their participation and the expected benefits for them were explained to the teachers. Full cooperation was granted in every case. The major hesitation centred around the video recording of the classroom. The teachers understood that this was essential for the research and that the recordings and information derived from them would be used confidentially with only those directly associated with the research having access. It was agreed that the recordings would be obtained with as little disruption as possible to their normal routine.

Early in the 1977 school year, the three schools were visited again to meet the staff involved, to ascertain when the Progressive Achievement Tests would be administered, and to obtain official school approval so that the Wellington Education Board could be officially approached. At this time, the principal of one school said that his school was in the Wanganui Education Board which has different procedures for conducting research. He had assumed that the experimenter had approached the Wanganui Education Board and had received approval. The experimenter had assumed because of the school's location and that it had been suggested by a Wellington Education Board employee that it was a school of that district.

The forms for applying to do research in the Wanganui Education Board were filled out and after two months of supplying additional justifications for the research proposal, the Wanganui Education Board granted permission to proceed with the investigation with the stipulation that the Progressive Achievement Tests not be used as a measure of student achievement. While there are many measures of educational achievement that are produced throughout the world, the Progressive Achievement Test series have gained justifiable regard as the most accurate and reliable group tests available that give an indication of achievement toward meeting some of the goals of New Zealand education. To place the restriction of not using these tests on the planned research would weaken the possibility of drawing valid conclusions on

several of the major hypotheses. As the Wellington Education Board had given their support and encouragement to proceed in all matters, another school in the Wellington Education Board was searched for, found, approached, and upon very short notice, video-taped.

All three schools were video-taped in the same manner. During a time when no students were in the classroom, two television cameras with remote control drives were placed on tripods in corners of the room where they would not be in the way but would still maintain a good coverage of the classroom. The lenses could focus, zoom, and control the amount of light admitted remotely. Two microphones were hung from the ceiling to record student talk. This was to be fed into one audio channel. The teacher wore a small microphone with a small FM radio transmitter. This permitted complete freedom of motion. The radio signal from the teacher was picked up by a receiver in the classroom and was to be recorded on the second audio channel. Unfortunately, because of a technical problem, both audio channels were recorded on one channel. The teacher was then free to start his teaching when ready. All recording and camera controls were performed by the two man team, located outside of the classroom. Thirty minutes of a mathematics lesson and thirty minutes of an art lesson were recorded for each teacher. If the teacher wished, the class was permitted to have a look at the inside of the mobile recording studio. The recordings were completed during two weeks in the middle of April, 1977.

Training sessions began the first week of the second term. These were conducted after school time, one day a week for approximately an hour each session. The ten sessions were completed by the last week of the second term. Attendance was not always perfect. At the experimental school one teacher dropped out of school entirely. Another teacher was admitted to the hospital and confined to bed rest for the last half of the second term. She returned for some of the third term and completed all post measures. In the placebo school the senior teacher retired shortly after the initial video-taping. The replacement teacher declined the invitation to participate. Another teacher was frequently missing from the training sessions as heavy sport's obligations kept him involved. He was approached and encouraged to be present, which he agreed to do, but he was only present for three of the ten sessions. At the two treatment schools five teachers were thus able to be recorded twice and data collected from their students. This number was balanced with five teachers from the control school.

The final video-recording went according to schedule during the second and third weeks of November. The one change was that the only microphone used was the radio microphone of the teacher. This picked up the relevant student talk and teacher talk more satisfactorily. Unfortunately, the answer sheets for the mathematics examination were not included with the Progressive Achievement Test materials posted from Wellington. Two of the schools had sufficient answer sheets on hand to complete this examination during the last week of classes, however, one school did not do this examination. Neither did they complete the attitude questionnaire.

;

## DATA ANALYSIS

The research programme incorporated measures of teachers and their pupils taken at the beginning and at the end of the school year. First the initial teaching behaviour of the teachers will be analyzed followed by an examination of their final behaviour. Then the major hypothesis will be tested, does the attending and responding behaviour of the teacher make a difference in the performance of the student?

### Attending - Reliability

A summary tape was constructed of ten minute segments of each of the five teachers from one school. The segments were selected on the basis of active teacher participation with individual children.

Three graduate education students participated in a ten minute description and discussion of the attending scale. Then, they began viewing the recording, noting the index time of the incident and their judgement of the teacher's attending behaviour based on the scale. The first fifteen minutes of viewing were spent clarifying the application of the scale to the observed teaching behaviours. After this, they were able to watch and code the remainder of the tape. There was a ninety per cent agreement as to when an incident occurred. Individually seventy-four per cent, sixty-five per cent, and seventy-four per cent of their codings agreed with an expert rater. After only a minimum of training time, this standard was deemed sufficient to proceed.

The tapes were subsequently coded by raters who were checked for inter-rater reliability and the amount of agreement with an expert rater. There was an inter-rater agreement on ninety-four per cent of the ratings for one complete lesson by one teacher and eighty-one per cent agreement with the expert rater. All ratings were done independently.

There was sufficient inter-rater reliability to justify using the attending scale. However, overall the teachers were not stable in their mean attending behaviour across subject matter and over time. Pearson product moment coefficients were determined for

initial teacher mean attending scores in mathematics and art ( $r = -.17$ ), for initial and final teacher mean attending scores ( $r = .10$ ), and initial and final attending scores of the non-experimental teachers ( $r = .04$ ). The experimentally trained teachers had a coefficient of  $r = .88$  for their initial mean attending behaviour compared to their final mean attending behaviour. There was an average increase with this group of 0.20 on the attending scale. They improved with their ability to attend in a highly reliable manner compared with their initial attending behaviour. The remaining teachers varied unsystematically.

TABLE 7

Product Moment Coefficients of  
Teacher Attending Scores

---

All teachers initial mathematics attending with  
art attending  $r = -.17$

All teachers initial attending with final  
attending  $r = .10$

Non-experimental teachers initial attending with  
final attending  $r = .04$

Experimental teachers initial attending with final  
attending  $r = .88$

---

Most previous attempts to identify teaching behaviours that are associated with student gains have encountered similar inconsistent teacher behaviours. (Shavelson and Dempsey, 1976) The rating process and scales to measure teacher behaviours have been developed to achieve high reliability levels. However, teachers have varied their teaching behaviours, whether systematically or randomly, over time. This has weakened previous attempts to identify those teaching behaviours that do make a difference.

### Teacher Attending Behaviour

A complication of conducting a field experiment in a naturalistic setting is the greater likelihood of having an unbalanced factorial design. The analysis of variance becomes somewhat more involved than it is in an orthogonal design where every cell has an equal number of cases. One noticeable effect of an unbalanced factorial design is that the component sum of squares frequently will not be equal to the total sum of squares. Thus, the combined effect of two variables can differ from the contribution of the variables working independently. This effect is accentuated when the two variables are highly correlated. It is possible for the additive effect of two variables to be significant while neither of the individual variables is significant. In the summary table presentation the sum of squares that is shown for a variable is that variable's total contribution of variance adjusted for the influence of the other variables at the same level. (Nie, et al, 1975)

From the initial video recordings there were no significant differences between the attending behaviour of the teachers at the three schools, between male and female teachers, nor between teachers at different levels. Early in the school year, the teachers tended to give more attention to the pupils during art. While this difference was significant, in absolute terms the average difference in attending was 0.12 units on the attending scale. There were no significant interaction effects.

TABLE 8

#### Analysis of Variance Teachers Attending Pre-test

Source of Variation	Sum of Squares	df	Mean Square	F Ratio	Level of Significance
Main Effects	0.165	4	0.041	1.52	0.24
Schools	0.032	2	0.016	0.59	
Teacher Sex	0.014	1	0.014	0.52	
Subject	0.124	1	0.124	4.55	.04
Two Way Interactions	0.082	5	0.016	0.60	
School x Sex	0.065	2	0.033	1.19	.32
School x Subject	0.013	2	0.007	0.24	
Sex x Subject	0.006	1	0.006	0.24	
Total Explained	0.247	9			
Residual	0.545	20	0.027		
Total	0.793	29			

The mean for all of the teachers' attending behaviour was 1.26. On the attending scale 1.0 is the absence of all major non-verbal indicators of paying attention, whereas, 1.5 is the addition of facing the person squarely. On average, these fifteen teachers were facing their children one half of the time while they were engaged in a one to one interaction. The other half of the time they were not giving any physical indication of paying personal attention. The range of teacher mean attending behaviour went from 1.08 to 1.44.

From the ratings of teacher attending behaviour taken at the end of the year there were no significant differences between the teachers of the three schools and between male and female teachers. There was a significant difference in teacher attending behaviour between mathematics and art lessons. Teachers gave more attention during mathematics lessons. The mean attending for mathematics was 1.50 and for art was 1.38. On the initial base line measures the teachers' mean attending behaviour for mathematics was 1.20 and for art was 1.33. During the course of the year the teachers did not change their attending behaviour during art. They did give more indications of attention at the end of the year during mathematics.

TABLE 9

## Analysis of Variance Teacher Attending Post-test

Source of Variation	Sum of Squares	df	Mean Square	F Ratio	Level of Significance
Main Effects	0.244	4	0.061	2.40	.08
Schools	0.097	2	0.049	1.92	.17
Teacher Sex	0.016	1	0.016	0.63	
Subject	0.123	1	0.123	4.84	.04
Two Way Interactions	0.300	5	0.060	2.36	.08
School x Sex	0.277	2	0.139	5.46	.01
School x Subject	0.017	2	0.009	0.34	
Sex x Subject	0.005	1	0.005	0.19	
Total Explained	0.544	9			
Residual	0.507	20	0.025		
Total	1.051	29			

There was a significant interaction of school and sex of teacher. The female teachers in the experimental school gained an average of 0.213. The female teachers in the placebo school went down -0.09 and the control school females went up 0.17. The males in the treatment school went up 0.135. However, the males in the placebo treatment increased 0.382, while the males in the control group increased by 0.09.

TABLE 10  
Increases in Attending by Teachers Sex

<u>School</u>	<u>Male</u> n	<u>Female</u> n	<u>Total Average</u>
Experimental	+0.135 (2)	+0.213 (3)	+0.182
Placebo	+0.382 (3)	-0.090 (2)	+0.166
Control	+0.090 (2)	+0.170 (3)	+0.138
Total average	+0.228	+0.121	+0.162

On average the teachers who received training to improve their attending behaviours did improve the most of the three schools. An average gain of less than 0.2 illustrates that these teachers were able to exhibit a higher level behaviour four times more out of ten behaviours than they were at the beginning of the year.

The control teachers gained familiarity with their students by being with them during the year. They exhibited a higher level of behaviour less than three additional times for every ten interactions. The F ratio of 1.92 shows that the probability of this difference happening by chance is less than .17. While this is not statistically significant, the difference is in the direction hypothesized. Training changed the classroom attending behaviours of the teachers.

### Responding - Reliability

In a review of the research methodology and the reporting of results of the Carkhuff training model Gormally and Hill (1974) encouraged the reporting of reliability levels obtained in all experiments and replications instead of depending on reliability levels reached previously. Sometimes there has been no reporting of reliability coefficients on scales similar to the one used in this research. Statements have been provided that the scale has been validated in process and outcome research and that "reliabilities were determined in previous studies to be in the .80's and .90's." (Pierce and Schauble, 1970; Gormally and Hill, 1974; Lambert and De Julio, 1977; Hefele and Harst, 1972)

Rogers (1967) in his process and outcome research with the counselling of schizophrenic patients developed raters until they were able to achieve a product moment correlation of  $r = .60$  for inter-rater reliability using written practice materials. Their actual reliability on the job was not calculated but only estimated by using the raters' mean scores.

In this project one rater was trained initially with role plays, examples, and didactic illustrations of each level of the Teacher Effectiveness Scale until there was agreement with an expert rater. On his own this rater then achieved  $r = .31$  with an expert rater. Another discussion and workshop ensued with subsequent ratings of the recording having a correlation of  $r = .36$  with an expert. The actual ratings were compared and it was found that 80.5 per cent of the ratings agreed exactly. Only three per cent of the ratings were more than one scale stop in difference. Checks after this time showed  $r = .50$  and later  $r = .54$  with high levels of agreement with an expert rater with only small differences.

Considering that the range of the verbal behaviour of the teachers was so limited as measured by this scale, from 1.0 to 1.5 to 2.0, the reliability coefficients were deemed satisfactory. Through increasing the range of teacher behaviour on the scale one can easily expect even higher reliability coefficients. (Wesman, 1952) In training sessions examples of the higher levels were discriminated accurately by the raters, however, in the classroom teachers were not utilizing these behaviours.

There was no correlation between all of the teachers overall pre- and post-measures of responding ( $r = .00$ ). Between teachers mean responding in mathematics and art there was a correlation of  $r = .35$ . Thus teachers responded at a relatively similar level on the same day in two subjects, but there was no relation to their responding level at the beginning and the end of the year. Interestingly, the five experimental teachers had a correlation of  $r = -.51$  of pre- and post-responding levels. On average the teachers who tended to be the most accurate at identifying the pupil's frame of reference in the beginning did not improve their level of responding, whereas the three teachers who initially were lower, passed them.

#### Teachers Responding Behaviour

From the initial video recordings there were no significant differences in verbal responding skills between the teachers of the three schools, between male and female teachers, nor between teachers during mathematics and art. Neither were there any significant interaction effects. Male teachers had a tendency to respond more accurately during mathematics while female teachers had a slight tendency to be more accurate during art. This pattern was not found at the end of the year.

TABLE 11  
Analysis of Variance Teacher Responding Pre-test

Source of Variation	Sum of Squares	df	Mean Square	F Ratio	Level of Significance
Main Effects	0.011	4	0.003	0.35	
Schools	0.007	2	0.004	0.46	
Teacher Sex	0.000	1	0.000	0.00	
Subject	0.004	1	0.004	0.48	
Two Way Interactions	0.051	5	0.010	1.28	.31
School x Sex	0.016	2	0.008	1.02	
School x Subject	0.012	2	0.006	0.77	
Sex x Subject	0.022	1	0.022	2.72	.11
Total Explained	0.063	9			
Residual	0.160	20	0.008		
Total	0.223	29			

The mean for all of the teachers' responding behaviour was 1.45. On the average these fifteen teachers responded to an individual talking with them by talking about the general subject of interest, but largely from the teachers' own frame of reference. The range of teacher mean responding behaviour was 1.32 to 1.57. Only once out of almost 1,500 teacher responses was the feeling of a pupil identified.

TABLE 12  
Analysis of Variance Teacher Responding Post-test

Source of Variation	Sum of Squares	df	Mean Squares	F Ratio	Level of Significance
Main Effects	0.099	4	0.025	3.23	.03
Schools	0.011	2	0.006	0.74	
Teacher Sex	0.002	1	0.002	0.31	
Subject	0.086	1	0.086	11.22	.003
Two Way Interactions	0.031	5	0.006	0.80	
School x Sex	0.001	2	0.001	0.09	
School x Subject	0.027	2	0.013	1.74	.20
Sex x Subject	0.002	1	0.002	0.29	
Total Explained	0.130	9			
Residual	0.154	20	0.008		
Total	0.284	29			

From the ratings of teacher responding behaviour taken at the end of the school year there were no significant differences between the teachers of the three schools and between male and female teachers. There was a difference in the level of responses between mathematics and art that was significant at the .003 level. In mathematics the teacher responding average was 1.60. While in art, it was 1.50. The teachers were slightly more accurate in demonstrating their understanding of the child's communication during mathematics. During the same period, the teachers also gave more attention in mathematics. (See Table 9.)

On average the teachers who received training to improve their responding behaviour did improve the most of the three schools, however, this difference was detectable only at the third decimal place.

TABLE 13

## Average Gain in Responding

Experimental	+0.108
Placebo	+0.106
Control	+0.062

Given that the responding scale identifies an additional skill for each 0.5 unit, the treatment school teachers are able to use an additional responding skill two additional times out of every ten interchanges.

The range of verbal responding is of interest as it was only in the experimental school that teachers were able to respond verbally to the feelings of the students. Of the slightly more than 800 teacher responses in the experimental school, thirty-two included accurate identification of the pupil's feeling. No other teacher in the other schools did this even once. In addition, of the five teachers in the experimental school, two teachers gave no feeling response at all and one teacher used only two responses that incorporated feelings. Possibly, it is only with training that these verbal responses will be achieved and conceivably ten hours of training may not be enough for the majority of teachers. One teacher employed a feeling response in 8.7% of her responses and the other teacher used a feeling response in 5.9% of her responses. Their overall average was not significantly different from all of the other teachers as they both had many lower level responses which levelled off their responding average.

The verbal and non-verbal teaching behaviours of the experimental school teachers improved in the direction desired, toward giving more indications that they are ready to receive information from their pupils and that they do understand where their pupils are. Almost all of the other teachers also improved along these dimensions, although not to the same extent. As groups there are no significant differences when the mean behaviour is analyzed. One would therefore expect that there will be no significant differences between the performance of the pupils that can be attributed to the mean attending and responding behaviour of the teachers.

A more meaningful comparison in that it is closer to testing the original hypothesis is to compare the students of those teachers who have demonstrated the highest level responding with the students of the teachers who have the lowest overall average and are teaching at the same class level.

The two teachers who were able to identify their pupils' feeling frequently were selected. Their average post-attending behaviour was 1.52 and 1.54. Their average post-responding behaviour was 1.63 and 1.59. These teachers were teaching in standards two, three, and four.

Table 14 lists the teachers at the same levels in the other schools, their averages, and a grand average taken by adding the attending and responding scores and obtaining the average.

TABLE 14  
Lower Primary Teachers' Average Scores

School	Attending	Responding	Grand Average
Experimental A	1.52	1.63	1.575
" B	1.54	1.59	1.565
Placebo A	1.82	1.55	1.685
" B	1.34	1.51	1.425
" C	1.27	1.54	1.405
Control A	1.44	1.52	1.480
" B	1.37	1.54	1.455
" C	1.36	1.48	1.420

Teacher "C" in the placebo school and teacher "C" in the control school were chosen to act as the greatest possible contrast to the two teachers selected.

The comparisons of their students' raw scores on the Progressive Achievement Tests of reading comprehension, vocabulary, and mathematics given at the end of the school year with their scores from the beginning of the school year acting as covariates are shown in Tables 15, 16 and 17.

TABLE 15

Analysis of Covariance: High Teacher vs Low Teacher -  
Reading Comprehension

Source of Variation	Sum of Squares	df	Mean Square	F Ratio	Level of Significance
Covariate					
Comprehension Pre-Test	2872.391	1	2872.391	99.09	.001
Main Effect					
Teacher High-Low	55.732	1	55.732	1.92	.17
Residual	2376.983	82	28.988		
Total	5305.106	84			

TABLE 16

Analysis of Covariance: High Teacher vs Low Teacher -  
Vocabulary

Source of Variation	Sum of Squares	df	Mean Square	F Ratio	Level of Significance
Covariate					
Vocabulary Pre-Test	8739.941	1	8739.941	241.373	.001
Main Effect					
Teacher High-Low	0.315	1	0.315	0.009	
Residual	2969.156	82	36.209		
Total	11709.412	84			

TABLE 17

Analysis of Covariance: High Teacher vs Low Teacher -  
Mathematics

Source of Variation	Sum of Squares	df	Mean Square	F Ratio	Level of Significance
Covariate					
Mathematics Pre-Test	3475.954	1	3475.954	70.10	.001
Main Effect					
Teacher High-Low	115.454	1	115.454	2.33	.13
Residual	2925.560	59	49.586		
Total	6516.968	61			

The difference in gains for reading comprehension does not reach statistical significance. However, on average the two experimental teachers' students score 1.74 more correct answers than the comparison teachers' students when adjusted by the pre-test scores.

The difference in the vocabulary test is much smaller, although it is in the hypothesized direction.

The difference in the levels of mathematics achievement is more noticeable. The F ratio of 2.33 is significant at the .13 level. While this is not strong enough to reject the null hypothesis, suspicions can be raised that a meaningful contrast has been raised especially when it is considered that on average the experimental teachers' students score 2.83 more correct answers than the comparison students when adjusted by the covariate.

Thus there are indications from these results that teachers who incorporate the feelings of their pupils into their responses tend to achieve some of the cognitive objectives more readily than teachers who do not give evidence of this understanding.

Whilst the sample is small, an indication of the strength of this factor in explaining academic achievement may be constructed through multiple regression onto the average raw score gain of the pupils of each teacher on the three Progressive Achievement Tests. Factors that will be regressed onto this average gain score of each teacher will be teacher sex, teacher class level, teacher experience, average post-

attending score, average post-responding score, school, and the number of teacher responses that incorporate pupil feeling in one hour.

The hypothesis is that the level of accurate responding behaviour of the teacher will be positively associated with student academic achievement. The variable of gross number of feeling responses appears to reflect this factor to a greater degree than does the average responding score. Because of this and because of the high intercorrelation of teacher sex and number of feeling responses and the further high correlation of both of these factors on pupil achievement, it was decided to enter teacher feeling responses into the regression first to obtain the maximum explanatory effect of teacher feeling responses on pupil achievement that this sample shows. This is an additional, direct test of the major hypothesis. In addition, this variable is a behaviour that any teacher can utilize, irrespective of such categorical variables of sex, level, school, and experience. These remaining variables and the average attending and responding scores will be entered into the regression on a stepwise progression with the variable that explains the largest amount of remaining variance of pupil achievement being entered next into the equation until there is no more significant amount of variation left.

TABLE 18  
Correlation Matrix

Correlation Matrix	Average Gain	Attend	Respond	Sex	Experience	Level	School	Feeling
Average Gains	1.00000	-0.02891	0.11296	0.74578	-0.26948	-0.35558	0.19909	0.40344
Attending		1.00000	0.05137	-0.19426	0.20122	-0.42511	-0.30177	0.22851
Responding			1.00000	0.08211	-0.11333	-0.14742	-0.29565	0.41529
Teacher Sex				1.00000	-0.09629	-0.18898	0.00000	0.30434
Teacher Experience					1.00000	0.11323	0.19612	-0.40040
Class Level						1.00000	0.00000	-0.39845
School							1.00000	-0.46009
Number of Feeling Responses								1.00000

Teacher Sex: Male = 1; Female = 2.  
 Teacher Experience: 1-3 years taught; 4-6 years; 7 or more.  
 Teacher Level: Standard 2 = 1; Standard 3 = 2; Standard 4 = 3; Form I = 4; Form II = 5.  
 Teacher School: Experimental = 1; Placebo = 2; Control = 3.

TABLE 19

## Multiple Regression onto Average Student Gain

Feeling responses entered first into the regression. The remainder entered stepwise.

Variable	Multiple R	R Square	RSQ Change	Simple R	F Ratio	B	Beta
Number of Feeling Responses	0.40344	0.16276	0.16276	0.40344	4.71	0.06756	0.21887
Teacher Sex	0.76845	0.59051	0.42775	0.74578	12.39	2.44835	0.69683
School	0.83614	0.69913	0.10861	0.19909	3.15	0.92968	0.43305
Teacher Experience	0.84782	0.71881	0.01968	-0.26948	1.00	-0.50612	-0.24042
Attending	0.86950	0.75603	0.03722	-0.02891	1.07	3.01558	0.23274
Responding	0.87082	0.75832	0.00230	0.11296	1.00	1.54852	0.05367
Constant						-6.09903	

The variable of teacher level was not entered into the regression as all of the variance that it accounted for was explained previously.

The number of teacher feeling responses accounted for sixteen per cent of the variance in pupil achievement. The more teachers were able to accurately identify their pupils' feelings, the more the pupils tended to achieve. Nie, et al (1975) offer a test of significance for the amount of  $R^2$  change.

$$F = \frac{R^2 \text{ change of the variable} / 1}{1 - \text{total } R^2 \text{ change} / df}$$

The degrees of freedom are one and the number of cases, minus the number of independent variables, minus one. For this regression the  $df = 1, 7$ .

The F ratio of 4.71 for teacher feeling response is significant between .10 and .05. Therefore, the probability of teacher feeling responses not affecting student achievement by chance as shown by the present data is less than .10 and greater than .05. While the null hypothesis usually has the criterion of .05 set for rejection arbitrarily, the null hypothesis still must be entertained that teacher feeling responses make no difference on pupil achievement. However, the null hypothesis has not been accepted as the true statement of reality. More investigation is needed, especially in broadening the sample of teachers who can use this skill in the classroom and under laboratory settings to further test the effects.

The variable of teacher sex explained forty-three per cent of the remaining variance in pupil achievement. The F ratio of 12.39 with 1 and 7 degrees of freedom is significant beyond the .01 level. For this sample we can confidently reject the null hypothesis that teacher sex made no difference. By inspecting pupil achievement gains and teacher sex in Table 20 it is shown that the pupils of female teachers had higher average gains.

What did these teachers do that was important? The correlation ( $r = .30$ ) of teacher feeling and teacher sex offers the possibility of one dimension of the difference. The female teachers in the experimental school included more of the responses that the training sessions hoped to evoke.

TABLE 20

## Rank Ordering of Teachers According to Average Pupil Gain

Average Pupil Gain	Sex	Level	Experience	Number of Feeling Responses	Attending	(% Responses With Eye Contact	Responding
8.16	F	1	1	19	1.52	(26.6%)	1.63
7.06	F	4	2	0	1.30	(12.4%)	1.67
6.86	F	2	1	0	1.34	(12.3%)	1.51
6.63	F	1	3	0	1.44	(21.2%)	1.52
6.60	F	5	3	0	1.34	(17.2%)	1.44
6.08	M	3	1	0	1.37	(14.7%)	1.52
5.83	F	3	1	0	1.27	(11.2%)	1.54
5.59	F	4	3	0	1.52	(23.4%)	1.50
5.10	F	2	2	11	1.54	(27.5%)	1.59
4.78	M	1	3	0	1.82	(33.6%)	1.52
4.52	M	4	2	0	1.56	(23.6%)	1.60
4.11	M	5	2	2	1.37	(22.2%)	1.51
3.49	M	2	3	0	1.36	(17.9%)	1.48
2.63	M	3	3	0	1.36	(15.2%)	1.62
1.39	M	5	2	0	1.39	(21.1%)	1.53
						(20.0%)	

Whether it is the actual response with the student or a broader teaching and interpersonal strategy of acquiescing or accepting the other person's point of view as a starting point in the learning process remains to be investigated. One of the principles of effective teaching has been to start where the learner is. Possibly, the female teachers are doing this more effectively. Two features do stand out: (1) female teachers over all schools and levels accomplished higher achievement; (2) in the experimental school the female teachers included more of the verbal responses that were one of the objectives of the training sessions.

The next most powerful predictor of pupil achievement was the school itself. Almost eleven per cent of the variance was explained by this factor. The F ratio of 3.15 was less than that needed to be significant at the .10 level. What was the overall factor that just being at one school instead of another tended to affect pupil performance? The contributing factors may be the socio-economic status of the parents of the children. One expects that if one school has parents of a higher socio-economic status, that their children will achieve slightly more even when entering achievement levels are accounted for.

The children in the upper three levels in the control school, which had the highest overall gain average, and the experimental school children had been asked what work their father and/or mother did. Only the upper levels were used for analysis as over generalization and ambiguity was more noticeable in the responses from children in standards two and three. The results were then coded based on the Socio-Economic Index for New Zealand developed by Elley and Irving (1972). This index ranks the highest levels based on education and income as 1 and the lowest level as 6.

TABLE 21

## Students' Parents Socio-Economic Status

	<u>1</u>	n	<u>2</u>	n	<u>3</u>	n	<u>4</u>	n	<u>5</u>	n	<u>6</u>	n
Experimental	6.9%	6	26.4%	23	12.6%	11	19.5%	17	19.5%	17	14.9%	13
Control	3.8%	3	40.5%	32	11.4%	9	21.5%	17	11.4%	9	11.4%	9

The frequencies of level one and two for each school were added together as were levels five and six.

- (1) Totals show that the control school had a greater proportion in the higher categories and a lesser proportion in the lower categories.
- (2) The chi square goodness of fit was used to determine if the frequency of higher level and lower level socio-economic conditions at the experimental school matched frequencies at the control school.
- (3) The chi squared statistic with the Yates' correction for continuity was 2.60.

The probability for such a balance to occur by chance would be approximately .11. Thus, the control school which had the highest overall gain score average also had higher socio-economic status of parents. Interestingly, the class from the experimental school with the teacher that used feelings frequently had a distribution that was predominantly (70%) lower class. Yet this class achieved more than any other class as measured by the average gain score.

The remaining four variables that were entered into the regression contributed smaller amounts of predicting power. Teacher experience contributed almost two per cent of the remaining variance. The less experienced teachers tended to have pupils that achieved at a higher rate.

The two averages of teacher classroom behaviour, post-attending and post-responding scores, contributed an additional four per cent of the variance. The F ratios for these two as well as teacher experience were less than one or just slightly more than one. They did not approach significance.

The level of the class contributed nothing in addition to explaining the gain score of the students.

The results from the open-ended attitude measure are equivocal as only fourteen out of twenty-five students from the target class in the control school replied and these were returned by mail after school break-up. Also the female teacher in the experimental school who was pregnant received a number of comments that she would have a boy or a girl. As the sex of the respondent was not known, no

inference was made whether this was a positive or negative attitude.

TABLE 22

## Pupil Attitude of Teacher

School	Positive	Neutral	Negative
Experimental	75%	22%	3%
Contrast	85%	7%	7%

## DISCUSSION AND CONCLUSIONS

Teachers aim at increasing the knowledge and the skills of their pupils. They have been trained to approach teaching with a frontal assault on performance. "Can the child read? Does he know his multiplication tables?" If he does not know seven times eight, he is told to learn seven times eight.

Evidence of this is in the performance of teachers as analyzed by their attending behaviour which showed that teachers give their prime attention to books, written work, and projects eighty per cent of the time. (Table 20) In only twenty per cent of the individual interchanges between a teacher and a pupil did the two have eye contact. Frequently, it was evident that the pupil followed the teacher's lead. If the teacher looked at the child, the child reciprocated. If the teacher focused on a thing, so did the child. The teacher with the highest percentage of maintaining eye contact did so in thirty-three per cent of the interchanges. The lowest level of eye contact was eleven per cent. Books, paper, and objects are important carriers of information and are proof of performance, but if the child is the centre and the purpose of education, it would seem logical to give him more attention.

Teachers are apparently available and approachable. In the average hour of videotaping, each teacher had 150 instances of paying attention to an individual. This is two and one half different interchanges per minute usually with different pupils. There are rapid conversations, rapid diagnoses, rapid solutions and rapid changeovers. Through it all, these teachers focus on things.

This is borne out still more when teacher response to pupil talk is analyzed. The verbal behaviour of teachers who were not trained in responding skills is largely restricted to ignoring or criticising the pupil, rewarding the child for agreeing with the teacher's answer, talking about the general area of pupil interest, and infrequently, identifying the specific content area of the child's concern. Only once out of 2,757 responses was the feeling of a pupil identified. Never was a deficit of a pupil mutually identified so that they could work on it.

There are indications in the present research that if the process

inherent in the teacher effectiveness scale is embarked on, gains in learning will be forthcoming. It is conceivable that equipping teachers in New Zealand with more of these humanistic-behaviouristic skills would produce results such as those seen by Aspy and Hadlock (Carkhuff, 1969) that so impressed Patterson (1973) and Good, Biddle and Brophy (1975).

From this sample of fifteen teachers the pupils of female teachers showed higher gains than pupils of male teachers irrespective of class level. There were male and female teachers at every level with only a tendency ( $r = -.18$ ) for females to be with lower class levels and males to be with the upper levels. There was a very strong correlation of female teachers with higher average pupil gains ( $r = .74$ ). Part of this can be explained by differences in classroom behaviours and even differences at the intake level of male and female applicants at teachers colleges in years past. Another contributing factor may be the perceptions of the children.

Webster (1973) found that male teachers college students were female oriented for support and preferred the emotional security of a safe institution. In contrast the female students exhibited more personal strength and self-confidence. Whilst these teachers college students differ from the usual expectation that maleness is equal to hardness and initiative and femaleness is equal to softness and responsiveness, the findings do indicate that the quality and actual behaviours of male and female teachers differ from each other and that part of this difference is operating at the time of selection into teacher training.

Children tend to see males as more powerful, more dominant, and more inclined to punish. Thus children generally tend to fear males more than females. (Jersild, 1968) This research found that female teachers tended to be more responsive to children and that their pupils achieved more than pupils of male teachers. If males enter teachers college with an inclination to docility and if they are not trained in responsiveness, then the results indicate that male teachers are deficient in both responding to pupils and in being able to act to direct pupil learning effectively.

To adequately test the original hypotheses will require more effective training with additional teachers as they are not naturally responding accurately to their pupils. Whether the breakdown in communication is

poor listening by the teacher, inadequate observing skills by the teacher, lack of attending by the teacher, or internal and external distractions that the teacher has no ability to control remains to be seen.

The necessity of training raises problems of effectively changing teacher classroom behaviour. In the small sample of five teachers who participated in the experimental training, two comfortably used the skill of responding accurately to pupil feelings. One other teacher did it twice in one hour in a pejorative manner and two did not do it at all. For these five teachers there was a correlation of  $r = -.81$  of the number of times feelings were identified in the classroom and years of teaching experience. One would not wish categorically to shape future training inputs on the basis of this negative correlation of length of service with empathy. The high correlation though does give one possible area of concern if training is to be widespread.

This finding may have a parallel in research done with the possibility of closed-mindedness of teachers. Heikkinen (1975) analyzed new entrants into a counselling training course and found that those teachers who were older tended to be more closed-minded which was thought to be detrimental to the learning of helping skills. Selvage (1977) attempted a replication and found no consistent pattern of bias. These results leave open the possibility that closed-mindedness is present among more experienced teachers and has a detrimental effect. While the behaviours of effective interpersonal communication are readily observable skills, their acquisition may be impeded if the trainee is not open to trying these new behaviours. One of the hindrances would be a dogmatic personality trait which may or may not be associated with age. Whatever the case, in order to enable all teachers to gain the skills, training will have to be conducted in a non-threatening environment by trainers who exhibit and model the desired behaviours consistently.

Ten hours of training in one hour units, separated by a week were not as productive as desired. Possibly, more extended blocks of time with less time in between sessions would have allowed more progress to be made. Subsequent to the research design and execution of the project, Carkhuff and Pierce (1976) published guidelines for the hours of training necessary to accomplish different levels of interpersonal skills. Their findings support the recommendation that more time for training

is necessary. They suggest 15-20 hours of training for a survey of interpersonal skills, 45-50 hours of training for the acquisition of the skills appropriate for brief interpersonal contacts, and 90 hours of training for the mastery of the skills necessary for on-going relationships.

As can be seen from tables 10 and 13, the placebo school teachers improved in attending and responding more than the control school teachers, yet not as much as the experimental school teachers. The latter group were explicitly taught these skills which were also modelled by the experimenter. The teachers at the placebo school were taught another aspect of educational information by the experimenter who at the same time modelled high level attending and responding skills. From this observation alone they were able to move in the hypothesized direction.

There appears to be support for the model of effective teaching presented here, but the evidence is not conclusive. The positive interaction of the affective and cognitive dimensions in the classroom is becoming more evident. While at one time it was popular to ignore emotions as detractors from the pursuit of truth, now it is more readily seen that feelings are integrally involved in the learning process. The teacher who is able to work with feelings will be able to facilitate more learning of subject matter. One would also suppose that this procedure could improve the human relations that exist in the classroom. This thesis did not investigate the affective benefits save through the attempt at attitude measurement. The specific validation of affective benefits could be a broader sphere for future investigation.

This study has been small in terms of the number of teachers who achieved part of the target behaviours in the classroom. The results indicate that a broadening of the number of teachers with the skills of attending and responding will make it possible for more generalizable conclusions to be reached. Further in-service training efforts are warranted.

The broadening of the sample of those who undergo training will provide increased data from classroom observation of teaching behaviour. This will further enable a testing of teacher characteristics and training procedures that lead to the acquisition or non-acquisition of relevant teaching skills. If the present finding of a very high negative correlation of teaching experience and the demonstration of target skills persists, one may question the most appropriate application of training. Pre-selection may be indicated and/or even pre-service

training in interpersonal skills. Indeed, it is logical to expect that continuous and intensive training opportunities will have to be available for teachers to develop and maintain their skills. This would help to prevent the loss of responsiveness over the years. There is also a responsibility to deliver benefits when the resources for training are limited.

There remain some of the original questions that have not been tested and indeed would be modified after this experience. Instead of testing the effectiveness of the complete model, are there critical steps that produce more benefits than others or is there a steady slope of increased benefits in terms of learning and emotional development? Or is there the possibility of greatly increased benefits after a certain stage is reached with only small benefits up to that stage? Information of this nature cannot be readily achieved in a classroom setting and points more to controlled, laboratory settings. This information can then be balanced against the training needs. To achieve the beginning stages of the model may require then to twenty hours of training. Whereas, the regular achievement of the entire process in the classroom may necessitate one hundred hours or more of intensive training if the training of teachers in these skills is parallel to the training of paraprofessionals in psychiatric work. (Carkhuff, 1969)

The results point to the need for further explication of the amount of pupil learning, both cognitive and affective, attributable to the level of teacher conditions offered. Such a project would also require careful re-examination of the amount of training necessary to achieve the various levels of teacher conditions. The teacher performance on initial video-taping and in the control and the placebo school at the conclusion strongly indicate that the delivery of the desired skills will not happen by chance. Specific, even systematic concentration upon the classroom delivery of relevant behaviours is essential in future training efforts.

The principles of good education have not been changed. The major contribution of the present study has been the identification of some of the neglected behaviours of teachers which appear to influence teacher effectiveness.

## BIBLIOGRAPHY

- Aspy, David (1972). Toward a Technology for Humanizing Education. Champaign, Illinois:Research Press.
- Aspy, David and Hutson, Barbara (1972). "Promotion of Student Success." Jnl. Educ. Res. 66:2:57-60.
- Berenson, David (1971). "The Effects of Systematic Human Relations Training upon Classroom Performance of Elementary School Teachers." Jnl. Res. and Devel. in Educ. 4:2:70-85.
- Bloom, Benjamin, et al (eds.) (1956). Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook I: Cognitive Domain. New York:David McKay.
- Brookover, Wilbur, et al (1964). "Self-Concept of Ability and School Achievement." Soc. of Educ. 37:271-278.
- Brophy, Jere and Good, T. (1974). Teacher-Student Relationships: Causes and Consequences. New York: Holt, Rinehart, & Winston.
- Brown, B., et al (1968). The Florida Taxonomy of Cognitive Behaviours. The Institute for Development of Human Resources. Gainesville:University of Florida.
- Brown, G. (1971). Human Teaching for Human Learning: An Introduction to Confluent Education. Auckland:Penguin.
- Campbell, D. and Stanley, J. (1963). Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally.
- Carkhuff, Robert (1969). Helping and Human Relations. Vol.I: Selection and Training. Vol. II: Practice and Research. New York:Holt, Rinehart, and Winston.
- Carkhuff, Robert (1972). "The Development of Systematic Human Resource Development Models." Counselling Psychologist. 3:3:4-11.
- Carkhuff, Robert and Berenson, Bernard (1967). Beyond Counselling and Therapy. New York:Holt, Rinehart, and Winston.
- Carkhuff, Robert and Berenson, Bernard (1976). Teaching as Treatment. Amherst:Human Resource Development Press.

- Carkhuff, Robert and Pierce, Richard (1975). Trainer's Guide: The Art of Helping. Amherst: Human Resource Development Press.
- Carkhuff, Robert and Pierce, Richard (1976). Teacher as Person. Washington, D.C.:National Education Association.
- Combs, Arthur (1965). "What Can Man Become?" In Hamachek, Don (ed.). The Self in Growth, Teaching, and Learning. Englewood Cliffs, New Jersey:Prentice Hall.
- Davidson, Helen and Lang, Gerhard (1960). "Children's Perceptions of the Teachers' Feelings toward Them Related to Self-perception, School Achievement, and Behaviour." Jnl. Exp. Educ. 29:107-118.
- Davis, O. and Tinsley, D. (1968). "Cognitive objectives revealed by classroom questions asked by social studies student teachers." In Hyman, R. (ed.) Teaching:Vantage Points for Study. Philadelphia: J.B. Lippincott.
- Dawson, R. and Hallinan, P. (1969). "Culturally Disadvantaged Children: A Study of Two Schools." Delta. 4:3-13.
- Day, H. and Berlyne, D. (1971). "Motivation and Learning." In Lesser, G. (ed.). Psychology and Educational Practice. Glenview, Illinois: Scott, Foresman.
- Dechant, Emerald (1968). Diagnosis and Remediation of Reading Disability. New York:Parker.
- Dunkin, Michael and Biddle, Bruce (1974). The Study of Teaching. New York: Holt, Rinehart, and Winston.
- Educational Testing Service (no date). Making Your Own Tests. Princeton, New Jersey:Educational Testing Service.
- Elley, W. and Irving, J. (1972). "A Socio-Economic Index for New Zealand Based on Levels of Education and Income from the 1966 Census." N.Z. Jnl. Educ. Stud. 7:2:153-167.
- Ellis, Albert (1962). Reason and Emotion in Psychotherapy. New York:Lyle Stuart.
- Eysenck, Hans (1960). "Learning Theory and Behaviour Therapy." In Behaviour Therapy and the Neuroses. London:Pergamon Press.
- Eysenck, Hans (1965). "The Effects of Psychotherapy." International Jnl. of Psychiatry. 1:99-178.

- Fink, Martin (1965). "Self-concept as it Relates to Academic Underachievement." In Hamachek, Don (ed.). The Self in Growth, Teaching, and Learning. Englewood Cliffs, New Jersey:Prentice Hall.
- Flanders, Ned (1960). Teacher Influence, Pupil Attitudes, and Achievement. Minneapolis:University of Minnesota.
- Flanders, Ned (1970). Analyzing Teaching Behaviour. Reading, Massachusetts:Addison Wesley.
- Gagné, Robert (1970). The Conditions of Learning, second edition. London:Holt, Rinehart, and Winston.
- Genthrer, Robert and Moughan, J. (1977). "Introverts' and Extroverts' Response to Nonverbal Attending Behaviour." Jnl. Counsel. Psy. 24:2:144-146.
- Gladstein, Gerald (1974). "Nonverbal Communication and Counselling/ Psychotherapy: A Review." Counselling Psychologist. 4:3:34-57.
- Glass, Gene and Stanley, Julian (1970). Statistical Methods in Education and Psychology. Englewood Cliffs, New Jersey: Prentice Hall.
- Gormally, James and Hill, Clara (1974). "Guidelines for Research on Carkhuff's Training Model." Jnl. Counsel. Psy. 21:6:539-547.
- Haase, R. and Tepper, D. (1972). "Nonverbal Components of Empathic Communication." Jnl. Counsel. Psy. 19:5:417-424.
- Harrison, F. (1968). "Relationship Between Home Background, School Success, and Adolescent Attitudes." Merrill-Palmer Quarterly of Behaviour Development. 14:331-344.
- Hebb, D. (1955). "Drives and the Conceptual Nervous System." Psy. Rev. 62:244 ff.
- Hefele, Thomas (1971). "The Effects of Systematic Human Relations Training upon Student Achievement." Jnl. Res. and Devel. in Educ. 4:2:52-69.
- Hefele, Thomas and Hurst, M. (1972). "Interpersonal Skill Measurement: Prediction, Validity, and Utility." Counselling Psychologist. 3:62-69.
- Heikkinen, Charles (1975). "Another Look at Teaching Experience and Closed-Mindedness." Jnl. Counsel. Psy. 22:1:79-83.

- Hess, R. and Shipman, V. (1971). "Maternal Influences on Early Learning." In Personality Growth and Learning. London: Longmans for Open University.
- Humphreys, Lloyd (1975). "Race and Sex Differences and their Implications for Educational and Occupational Equality." In Maehr, M. and Stallings, W. (eds.). Culture, Child, and School. Monterey, California: Brooks/Cole.
- Hunt, J. McV. (1968). "Experience and the Development of Motivation: Some Reinterpretations." In Hamachek, Don (ed.). Human Dynamics in Psychology and Education. Englewood Cliffs, New Jersey: Prentice Hall.
- Jersild, Arthur (1951). "Self-understanding in Childhood and Adolescence." Amer. Psych. 6:122-126.
- Jersild, Arthur (1955). When Teachers Face Themselves. New York: Columbia University.
- Kerlinger, Fred (1964). Foundations of Behavioural Research. London: Holt, Rinehart, and Winston.
- Kerlinger, Fred and Pedhazur, Elazar (1973). Multiple Regression in Behavioural Research. New York: Holt, Rinehart, and Winston.
- Krathwohl, David, et al. (1964). Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook II: Affective Domain. New York: David McKay.
- Kratochvil, D. (1968). The Effect of Parent and Teacher offered Levels of Facilitative Conditions upon Indexes of Student Physical, Emotional, and Intellectual Functioning. Unpublished Ph.D. Thesis, State University of New York at Buffalo.
- Krumboltz, John (1966). "Promoting Adaptive Behaviour." In Krumboltz, John (ed.). Revolution in Counselling: Implications of Behavioural Science. Boston: Houghton Mifflin.
- Lambert, Michael and DeJulio, Steven (1977). "Outcome Research in Carkhuff's Human Resource Development Programmes: Where is the Donut?" Counselling Psychologist. 6:4:79-86.
- Levitt, E. (1963). "Psychotherapy with Children: A Further Evaluation." Behaviour Research and Therapy. 1:45-51.
- Lindquist, E.F. (1940). Statistical Analysis in Educational Research. Boston: Houghton Mifflin.

- Lindquist, E.F. (1953). Design and Analysis of Experiments in Psychology and Education. Boston: Houghton Mifflin.
- Lumsdaine, A. (1962). "Instruments and Media of Instruction." In Gage, N. (ed.). Handbook of Research on Teaching. Chicago: Rand McNally.
- McClelland, David (1965). "Toward a Theory of Motive Acquisition." Amer. Psych. 20:321-333.
- Mehrabian, A. (1972). Nonverbal Communication. Chicago: Aldine-Atherton.
- Murray, Henry (1938). Explorations in Personality. New York: Oxford.
- Nie, (et al.) (1975). Statistical Package for the Social Sciences, Second Edition. New York: McGraw Hill.
- Patterson, Cecil (1973). Humanistic Education. Englewood Cliffs, New Jersey: Prentice-Hall.
- Pierce, R. and Schauble, P. (1970). "Graduate Training of Facilitative Counselors: The Effects of Individual Supervision." Jnl. Counsel. Psy. 17:210-215.
- Purkey, W. (1970). Self-Concept and School Achievement. Englewood Cliffs, New Jersey: Prentice-Hall.
- Ragosta, M. , et al. (1971). Sign versus Category: Two Instruments for Observing Level of Thinking. Paper presented at the annual meeting of the American Educational Research Association.
- Reed, H. (1961). "Effects of Teacher Warmth." Jnl. Tea. Educ. 12:330-334.
- Reed, H. (1961) "Teacher Variables of Warmth, Demand, and utilization of Intrinsic Motivation Related to Pupils' Science Interests." Jnl. Exp. Educ. 29:203-229.
- Reid, N. and Hughes, D. (1974). Progressive Achievement Test: Teacher's Manual: Mathematics. Wellington: New Zealand Council for Educational Research.
- Rogers, C. (ed.) (1967). The Therapeutic Relationship and Its Impact. Westport, Connecticut: Greenwood Press.
- Rogers, C. (1969). "Personal Thoughts on Teaching and Learning." In Rogers, C. Freedom to Learn. Columbus, Ohio: Charles E. Merrill.

- Rogers, C., et al. (1948). "The Role of Self-understanding in the Prediction of Behaviour." Jnl. of Consult. Psy. 12:174-186.
- Rogers, Virginia and Davis, O. (1970). Varying the Level of Cognitive Questions in the Classroom. Paper presented at the annual meeting of the American Educational Research Association.
- Selvage, Rob (1977) "Comment on Another look at Teaching Experience and Closed-mindedness." Jnl. Counsel. Psy. 24:2:142-143.
- Shavelson, R. and Dempsey-Atwood, N. (1976). "Generalizability of Measures of Teaching Behaviour." Rev. of Educ. Res. 46:4:553-611.
- Solomon, G. (1970). "The Analysis of Concrete to Abstract Classroom Patterns utilizing the TIP Profile." Jnl. of Res. and Devel. in Educ. 4:52-61.
- Solomon, G. and Wood, S. (1970). Classroom Behaviour Accompanying TIP Profile Measures. Unpublished, West Virginia University.
- Staines, J. (1958). "The Self-picture as a Factor in the Classroom." Brit. Jnl. of Educ. Psy. 28:97-111.
- Swift, D. (1971). "Social Class and Achievement Motivation." In Personality Growth and Learning. London: Longmans for Open University.
- Tatum, Carl (1972). "Humanism in Education." In Aspy, David. Toward a Technology for Humanizing Education. Champaign, Illinois: Research Press.
- Truax, Charles and Carkhuff, Robert (1967). Toward Effective Counselling and Psychotherapy: Training and Practice. Chicago: Aldine.
- Webster, A. (1972). Structure and Relationships of Standard Environmental Personality and Ability Factors in Secondary School Adolescents. Ph.D. Thesis. Palmerston North: Massey University.
- Webster, A. (1973). "Factors in the Effects of Education on the Human Resource." In Thomson, K. and Trlin, A. Contemporary New Zealand. Wellington: Hicks Smith.
- Wesman, A. (1952). Reliability and Confidence. Test Bulletin Number 44. New York: The Psychological Corporation.

- Wood, S. (1970). "A Multidimensional model for observation, analysis, and assessment of classroom behaviour." Jnl. of Res. and Devel. in Educ. 4:84-97.
- Wylie, Ruth (1974). The Self Concept: Revised Edition, Vol. I. Lincoln: University of Nebraska Press.
- Yablonsky, L. (1965). The Tunnel Back: Synanon. New York: Macmillan.