DOCUMENT RESUME

ED 193 256 TM 800 570

TITLE	Set Number One, 1980. Research Information for Teachers.
INSTITUTION	Australian Council for Educational Research, Hawthorn.; New Zealand Council for Educational Research, Wellington.
PUB DATE	80
NOTE	74p.: Items 6 and 7 are excluded because of copyright restrictions. For previous Set see ED 178 531.
AVAILABLE FROM	NCZER, Box 3237, Wellington, New Zealand ($10.00: 2 sets yearly)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	Classroom Environment: Cost Effectiveness: *Educational Assessment: Educational Environment: *Educational Improvement: Elementary Secondary Education: *Foreign Countries: Individualized Instruction: Instructional Innovation: Programed Instruction: Student Behavior: Teacher Attitudes: Teacher Made Tests: Textbook Evaluation
IDENTIFIERS	*New Zealand: Teacher Expectations

ABSTRACT

Topics and issues relating to educational research are presented in the form of separate news sheets or essays. A case study of the merits and deficiencies of Penrose High School's "Whanau House," is included as an example of curriculum innovation. An exploratory study describing expectations of New Zealand teachers demonstrates how teacher expectations cause certain students to be stereotyped and treated differently. An economist looks at what it costs society to invest in children's education, and what it pays. Some of the effects of school size on the behavior of pupils are reviewed. Two very different techniques to fit the work to the child (student team learning and self-paced instruction) are described. One article describes research findings on how successful Internal Assessment has been in Queensland's secondary school program. A summary of the techniques, theoretical basis, and changing nature of programmed learning is presented, as well as its impact on teaching. Other miscellaneous items of information include a guide to making up classroom tests, an analysis of textbook English, and a guide to class nomenclature. (RL)

set research information for teachers
number one 1980

set: research information for teachers, is published twice-yearly by the New Zealand Council for Educational Research and the Australian Council for Educational Research.

General Editor: Lyn Richards
Australian Editor: Peter Jeffery
Designer: Peter Ridder

New Zealand:
Subscriptions: $10 per year
Student subscription $8 per year
Discount subscriptions for bulk orders
To: **set**
NZCER,
Box 3237,
Wellington.

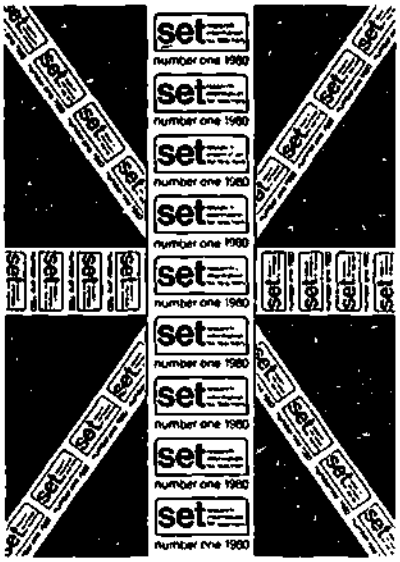
Australia:
Subscriptions: $10 per year
To: **set**
ACER,
Box 210,
Hawthorn,
Victoria 3122.

Item 1
(NZ)

Content Sheet

Research information for everyone interested in education: teachers, principals, students, lecturers, actively involved parents...

Leaflets and brief reports designed for private study, staff meetings, in-service courses, and discussion in small groups.



set
number one 1980

Produced twice-yearly by the New Zealand and Australian Councils for Educational Research

Subscriptions: $10 per year; Student subscription $8 per year: from:
set,
NZCER,
Box 3237,
Wellington.

Contents

1 Contents sheet

2 The Penrose Whanau Unit: A Case Study *Rae Munroe*
Like open-plan classrooms the Wanau units now being built demand more than just a change in where the children hang up their coats. An examination of hopes fulfilled and unfulfilled.

3 Teacher Expectations and Classroom Behaviour *Alison St George*
What do New Zealand teachers expect of their pupils? Are their expectations culturally stereotyped? Do they treat some pupils differently because of their stereotypes?

4 The Economics of Children in the Welfare State *Brian Easton*
An economist looks at what it costs society to invest in children and their education — and what it pays.

5 Projecting Educational Expenditure *Ian McGill and others*
What does it cost to educate one child? What will it cost in 1990? How will education cuts hurt? Are they necessary?

6 Chips With Everything excluded
Three articles about computers and education. A general account of where computers fit into schools; an article describing one successful use of a computer simulation in science; a light hearted account of the trials and tribulations of a Computer Assisted Instruction expert in Cincinnati.

7 How Many People Can a Young Child Feel Secure With? *Peter Smith* excluded
Is it safe, emotionally, to pass the baby from hand to hand, to leave it in child care, to take it to parties, to leave it in hospital? An article from Britain.

8 Small is Beautiful? *Jack Campbell*
Some of the effects on school size on the behaviour of pupils.

9 Class Size Rides Again
Does the size of a class have an effect on the achievement of the children in the class? The economic situation and new research techniques by *Gene Glass* and *Mary Lee Smith*, make a re-examination of the problem timely.

10 Catering for Individual Differences *Paul Power & John Longbottom*
Two very different techniques to fit the work to the child are described in this item: Student Team Learning, and an example of self-paced instruction

11 Assessing Internal Assessment *Ann Lamont and Bruce McBride*
Queensland has had Internal Assessment for several years now. This article describes research on how successful it has been.

12 What Ever Happened to Programmed Learning? *Graham Wagner*
Teaching Machines, Programmed Texts, etc., and their contribution to our understanding of how learning takes place.

13 Assessing What They've Learned *Warwick Elley*
A guide to making up classroom tests which are reliable and valid as well as quick and useful.

14 Educational English As She Is Written *Charles Taylor*
Textbooks as well as teachers put unnecessary obstacles in second-language learners' ways. Teachers can learn from this analysis of textbook English.

15 A Guide to Class Nomenclature
A 'already reckoner' chart to help with such mysteries as what an American book means by *Grade 3*, an Australian book by *Year 3*, and a New Zealand book by *Standard 3*.

~~Bonus~~ Item excluded

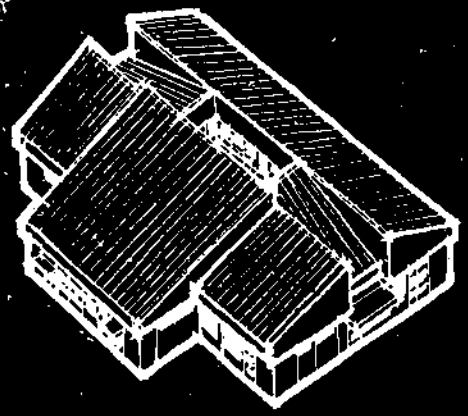
Free copies of *set* items

Copies of many past *set* items are available on request while they last. If the item is available its numerical list number is shown here. (See the *set* index, item 14 of *set* No. 1, 1979.)

63 80 86 97 103 113 116 120 123 127 132 138 141
73 81 87 98 105 114 118 121 124 130 133 139 142
77 83 93 99 109 115 122 125 131 134 140 143 148

Also available are copies of the following items from *set* No. 1, 1979: 3 6 7 8 10 11 12 14

and *set* No. 2, 1979: 4 5 6 8 9 10 11 12 13 14 15

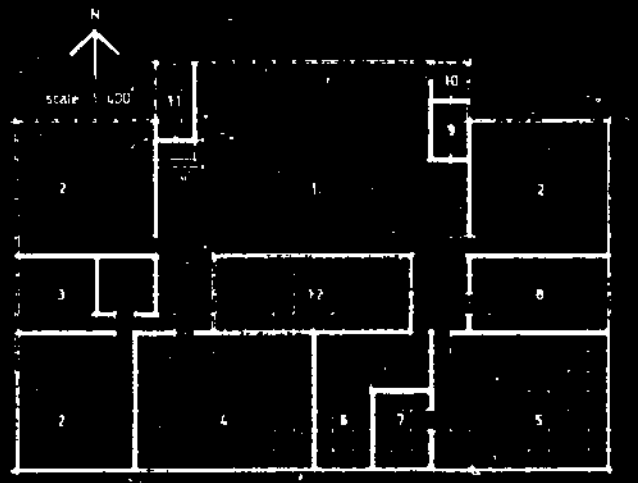


Block A : The Whanau House

The whanau house is entered via the Commons which is a large general purpose room. It also acts as a main circulation area and has provision for pupils' lockers. Associated with the entrance are the property office with its small refreshment counter, the Head of Department's office and a small interview room. On either side of the Commons area are classrooms, seminar rooms and a staff room. The long side of the Commons is bounded by a classroom, activities room, resources area, preparation area and a laboratory.

KEY

1	Commons Areas
2	Classroom
3	Seminar Rooms
4	Activities Room
5	Laboratory
6	Resources Room
7	Preparation Room
8	Staff Room
9	Head of Department
10	Interview Room
11	Property and Snack Bar
12	Internal court



The Penrose Whānau Unit

A Case Study

Whānau: 1. v.i. *Be born* I muri tonu mai tenei o tona whanautanga.
2. *Be in childbed* Ka rite nga ra e whanau ai te wahine.
3. n. *Offspring, family group* Ka rangona e Rala e karakia haere ana mai te whanau a Tane. Used occasionally in tribal designations, as, Te Whanau a Apanui.
4. *Family* (modern) It is questionable whether the Maori had any real conception of the family as a unit.
5. A familiar term of address to a number of people. Engari, e te whan. kei aroha tatou ki tona auetanga.

Pronunciation: as in, The lad will go *far now* as he has a good job.

Whanau: 1. v.i. Go.
2. *Lean, incline, bend down* Kia whanau tou taiepa – ka whanau iho te rakau.

Pronunciation: as in, The lad will have *fun now*.

The Penrose Whānau Unit

A Case Study

By Rae Munroe,
Auckland Secondary Teachers' College

Introduction

The original planning committee of the Whanau House design stated its expectations in *Secondary Schools for Tomorrow*. The expectations were that the House would be 'more intimate', provide 'room for spontaneity', be 'guidance centred', and 'give each pupil insights, knowledge and experience so that he can understand himself and the people he lives and works with, in both the smaller and the larger community; and to make the most of what he can do personally that is unique and vital to him', and further, that 'Dignity and self-esteem are essential to every human being, as is the feeling of belonging. Within the extended family of each Whanau House pupils will feel accepted and of value'.

The staff of the Whanau House built at Penrose High School has expectations too. These they wrote about in a booklet early in 1977. This statement was more clearly related to actual conditions in a New Zealand school and can be summarised as:

General Educational Expectations

learning would take place in a 'variety of ways and environments';
learning would be 'a *shared* experience with ... fellow classmates, teachers, parents and members of the community';
pupils' social and emotional needs would be met 'in a guidance-centred situation';
the Whanau would be a place 'to which the pupil belongs and with which he identifies'.

Expectations for Curriculum Organisation and Administration

Students should have a contribution to make to decisions concerning the curriculum and administration of the House.
Students should be allowed as much flexibility as possible in the choice of subjects including some reduction of the compulsory core and substitution of complementary courses so a wide range of topics may be offered.
Students should have the opportunity within subjects to involve themselves in remedial or extension programmes in addition to core requirements.
The facilities of the school and services of subjects should be available to the student outside the presently accepted school hours.

How many expectations were realised?

There were constraints (both anticipated and unanticipated) to the autonomy of a Whanau House staff, and some of these made it impossible to realise certain expectations. It would be unacceptable to judge the efforts of the staff by the number of expectations fulfilled. Penrose High School is a fairly typical urban secondary school, so it was inevitable that there would be a number of constraints on independence. The school had well established administrative routines and agreed expectations for the behaviour of pupils and teachers. The major constraints could be seen as associated with: the school timetable; the school's expectations of teachers; the school's expectations of pupils. Expectations only marginally affected were: learning, environment, sharing, guidance, and identity. Those affected directly were: curriculum and administrative decisions, curriculum choice, additional courses within the core, remedial and extension programmes, and extended hours.

Thus, the *General Educational Expectations* seemed the more realistic and most likely to be attained in practice. Those for *Curriculum Organisation and Administration*, although theoretically capable of realisation, were the most constrained by the existing conditions in Penrose High School.

Three main themes

The findings of the two-year case-study of the unit can be seen as having three main themes which stem from the expectations: The Whanau Unit as Community; Teaching and Learning in the Unit; and, The Unit as Democracy. The evidence upon which these commentaries are based comes from tape transcripts of interviews with pupils and teachers (inside and outside the unit), staff and staff-pupil meetings and seminars, classroom activities, and staff meetings with parents.

The Whanau Unit as Community

As soon as the teachers and the pupils entered the unit in April 1977 their comments reflected their belief that the Whanau House provided new opportunities for close association between pupils and between teachers and pupils. The comfortable conditions, the intermingling of teachers and pupils of all ages, and the relaxed and casual relationships were noted by all teachers and pupils during the intitial round of interviews.

A community developed and was seen in the extent to which pupils were 'looking after' their new environment. Care of equipment and furnishings was one constant perceived by all teachers and pupils during the two-year period of study. Although the standard of care varied somewhat from season to season, depending on the number of people using the building, teachers agreed that the pupils' attitude towards their environment was superior to that of pupils they had known in other school conditions. This opinion of Whanau staff members was frequently confirmed by visitors to the unit. On the whole teachers believed that the pupils' attitude was a function of their sense of identity with the building and with the people in it and this was confirmed by most pupils.

The range and the relaxed nature of friendship patterns among pupils were other aspects of community life commented on. However, sixth formers were reluctant to identify with the new unit and this reluctance persisted throughout the first year. In the second year the problem was less noticeable because the sixth formers had already had time to adjust to the unit.

Perhaps the most notable development in friendship patterns was the mixing of different age groups in the

commons. Teachers often contrasted the behaviour in the Whanau commons with their experience in the main school where 'juniors and seniors simply don't mix'. Nonetheless, teachers differed in their explanations of the change. Most attributed it, in part, to the building, but some saw the vertical form arrangement as a significant contributor to the mixing of age-levels, (yet this organisation was shared with the main school). The pupils themselves believed the mixing was wholly due to conditions in the commons 'where you naturally get to know everyone in the House'. In all, both teachers and pupils saw the socialising of pupils in the Whanau as significantly different from that in the main school. They perceived that barriers had been broken down for pupils both at their own age levels and between age levels.

The relationship between teachers and pupils was another sign of the Whanau House as a community. Inevitably teachers found that pupils had more access to them and although they all considered this to be a desirable trend there was sometimes an ambivalence in their comments. A few teachers, however, did not feel that the nature of their relationship with pupils had changed but simply that their own access to pupils was easier. Those teachers who did perceive marked changes in their involvement with pupils, nevertheless, had misgivings which in the initial stages were somewhat masked by the public commitment to persist with the goals established during planning. In the event, teachers, initially, made every effort to increase time spent in the unit while seeking to balance involvement against the desire to rest and be 'secluded'.

Despite high expectations for the development of a positive social climate in the early days and despite obvious gains in this area, by late in term three of the first year there were obvious signs in the teachers' comments that the experience had left them somewhat disillusioned. The changed emphasis can be traced in the transcripts of staff planning sessions. By November when the whole staff met for a day to assess progress, discussion tended to focus on some of the negative features of the year's experiences. In particular, the teachers' constant exposure to pupils had become an irritation to some of them. Given that the main thrust of the Whanau was to reduce the psychological distance between the teacher and taught, the reality proved to be too taxing for those teachers who were used to greater privacy in their day-to-day work at school. The lack of privacy in the staff planning centre led some staff to suggest that teachers in a Whanau needed less class contact time with pupils so they could cope with the additional demands on their time outside class hours.

In summary, the evidence suggested that for most pupils the unit promoted a very satisfying social environment which allowed for the establishment of wide-ranging friendships with peers and for increased access to teachers at both the personal and academic level. For some teachers the environment provided equally satisfying relationships, but for most the close social interaction resulted in varying degrees of fatigue and strain.

Teaching and Learning in the Unit

How did teachers perceive the effects of the move to the Whanau on classroom events? Their comments lay on a continuum from 'nothing's changed' to 'a completely new experience'. Some teachers had been able to use the facilities in the new building to advantage in extending a particular style of teaching. The extent to which this occurred varied between subjects and teachers but it was only a minority of the teachers who felt their teaching style had been markedly affected by the move.

At the same time there was a general consensus that the building was pleasant to work in, that this tended to make classroom activities more relaxed and that as a result pupils were better behaved. Again, the flexibility of the building was claimed as the main factor affecting change. The new environment, providing as it did carpeted areas, attractive decor, and easily moved furniture, was very much appreciated by those teachers who were promoting creative activities, group and project work, and informal discussion between pupils and themselves. However, a few teachers were concerned that flexibility in the unit had led to a certain slackness. Clearly the more relaxed routines did not suit all teachers equally. It was inevitable that those teachers who, for example, released pupils before the end of a period to go into the commons, were seen as eroding obedience to bell times. Other teachers who were more punctilious in this matter were upset by the apparent offhand approach of the pupils. 'Whanau pupils wandering about in corridors during class time' were frequently noted unfavourably by main-school teachers.

Thus, in the early months, the main teaching issues commented on by teachers related to the effects of the building and consequent changes in routines; examples were the fact that time was saved between periods because pupils and teachers did not have to 'walk down miles of corridors', the convenience of dividing classes between rooms, the ease with which teachers could find pupils, and the more mature behaviour of pupils. Some comment was made about curriculum development but largely in regard to the functioning of innovations which had been introduced before the unit was occupied.

For the individualised mathematics programme in the third and fourth forms the accommodation was thought to be ideal and the double activities room was successfully adapted for this work by the two mathematics teachers. The *Australian Science Education Project's* modules were introduced progressively in the junior science classrooms as the materials became available and interdisciplinary studies were begun for the third formers. Predictably, by the third term of the first year more detailed comment on the functioning of the various curriculum innovations was beginning to appear in interviews and in staff planning sessions.

Perhaps the single curriculum innovation which involved most planning, and could be said to typify the main thrust of innovation in the unit, was interdisciplinary studies. From the outset interdisciplinary studies was intended to embody three main intentions: to provide an opportunity for integration of core studies in junior classes, to allow for the full teaming of the Whanau teachers, and to give pupils opportunities for independent or self-directed learning. As well, the interdisciplinary studies were meant to alternate in blocks of three or four weeks with ordinary classwork in English, Social Studies, Science and Mathematics. Themes had been prepared at the planning stage and detailed work sheets with question sequences were designed before the studies were launched. The first session was held late in term two of the first year with third formers. Two further sessions, one third form and one fourth form, were held in the final term.

By the end of the first year there were differences between staff about the value of the studies, there was

an ambivalence in the response of the pupils, and a negative attitude to the studies by some parents and some staff members in the main school. These all contributed in some measure to a reduction of impetus for the innovation. As well, staff members found it increasingly difficult to find time to do the detailed planning necessary. In the second year only one third form session was held. By the end of that second year some teachers had reflected on the apparent failure of the innovation and commented on what they saw as shortcomings in practice.

The effect of the timetable structure was frequently mentioned as a major problem, and there is little doubt that it was a factor in the run down of interdisciplinary studies. So closely was the Whanau unit timetabled with the main school that any development such as this needed protracted negotiation, not only with main school teachers likely to be affected by Whanau pupils leaving the school during classtime for study in the community, but also for satisfying the main school's regulations for pupil's movements inside and outside the school.

Teachers perceptions of teaching and learning in the Whanau did centre on some common ground but there were also marked differences between their viewpoints. The teachers shared the opinion that they were in the presence of a more informal teaching environment than previously, that relationships with pupils had become less distant and that pupils had more ready access to them for special attention and tuition. However, opinion varied about whether or not pupils were learning more than they would have done in the main school and there were sharp differences in their perceptions of the extent to which curriculum change had actually taken place.

Comparison between 1978 School Certificate results of Whanau pupils and those of a matched group in the main school showed no significant difference in performance. However, Whanau pupils did pass, on average, a greater number of papers and this difference was statistically significant. On balance, it could be confidently claimed that the pupils who entered the Whanau unit were not academically disadvantaged at the fifth form level.

The Unit as Democracy

At the planning stage, participation by pupils was regarded as a major goal for the Whanau unit and early in the occupation of the building efforts were made to set up a House Committee which would encourage the pupils to participate in the affairs of the unit. However, the Whanau staff members had two clearly different positions about pupil-participation. One group accepted the original notion from the planning stage that format structures should be provided; the other group preferred the participation to emerge on a day-to-day basis built around teachers' contacts with individuals and small groups. Most discussions about participation were about the pupils involvement in the general running of the unit. However, a second aspect of participation, which had featured in the planning stage, was the pupils' contribution to curriculum development. This was little mentioned by teachers either in interviews or in meetings, and systematic feedback from pupils on curriculum matters was never a prominent consideration in the unit.

In all, the formal structures of vertical forms, House Committees, House meetings and forums (held infrequently, on Friday afternoons), were not seen by staff as fulfilling the high hopes held for them. The House meetings and vertical forms were seen to function successfully for administrative purposes in the school, the checking of absences, the communication of routine notices and for teachers to tell pupils about curriculum matters, field trips, school rules and the like. But they were not seen as providing many opportunities for pupils to participate either in the running of the unit or in the selecting and organising of the curriculum. Some teachers believed there were informal opportunities for pupils to participate in these ways but there was no consensus about the extent to which these opportunities existed for all pupils.

Conclusion

It is not possible to determine how much more could or should have been done to attain the goals of the Whanau. At the planning stage it was clear that the highest priority was placed on curriculum innovation. Innovation in consultation and decision-sharing, either between teachers and pupils or between teachers and parents, never engaged the full attention of the staff. Given the constraints noted, the situation that developed was not surprising.

There is no doubt that the Whanau staff worked extremely hard during the two years of this study and the range of their activities, both inside and outside the unit, was impressive. However, whether or not the goals would have been more successfully attained had there been additional staffing and therefore less contact-time for each teacher, is problematic. There is abundant evidence in the transcripts of the teachers' views that there was no clear consensus about which goals should have the greatest priority. Agreement was greatest about the social goals of the unit, but even here there were those who would have preferred less contact with pupils and greater protection of their privacy. In goals associated with teaching and learning, some teachers took and maintained an innovative stance while others did not. However, in the twin concerns of pupil-involvement and parent-involvement in the running of the unit there was considerable ambivalence in the views expressed and no systematic plan of action was adopted during the two-year period.

The course of events in the unit illustrated quite graphically the effects of unanticipated constraints on an educational innovation. A degree of autonomy had been assumed which could not be realised in practice. Without such necessary independence the Whanau staff had been presented with the impossible task of literally restructuring secondary schooling. The extent to which they succeeded, particularly in attaining the social goals, was clearly attributable to their commitment and their determination to persist in the face of considerable frustration.

Notes

Department of Education, *Secondary Schools for Tomorrow*, 1975.

Penrose High School, *Whanau House*, 1977.

Rae Munroe is Dean of the Secondary Teachers' College, Auckland and was Chairman of the PPTA Curriculum Review Group which produced the important book *Education in Change* in 1969.

The dictionary extract at the begining is from *A Dictionary of the Maori Language*, by H.W. Williams, Seventh Edition, The Government Printer, 1971, slightly shortened. The pronunciation guide is by the editor.

Item 15

set research information for teachers
number one 1980

A Guide to Class Nomenclature

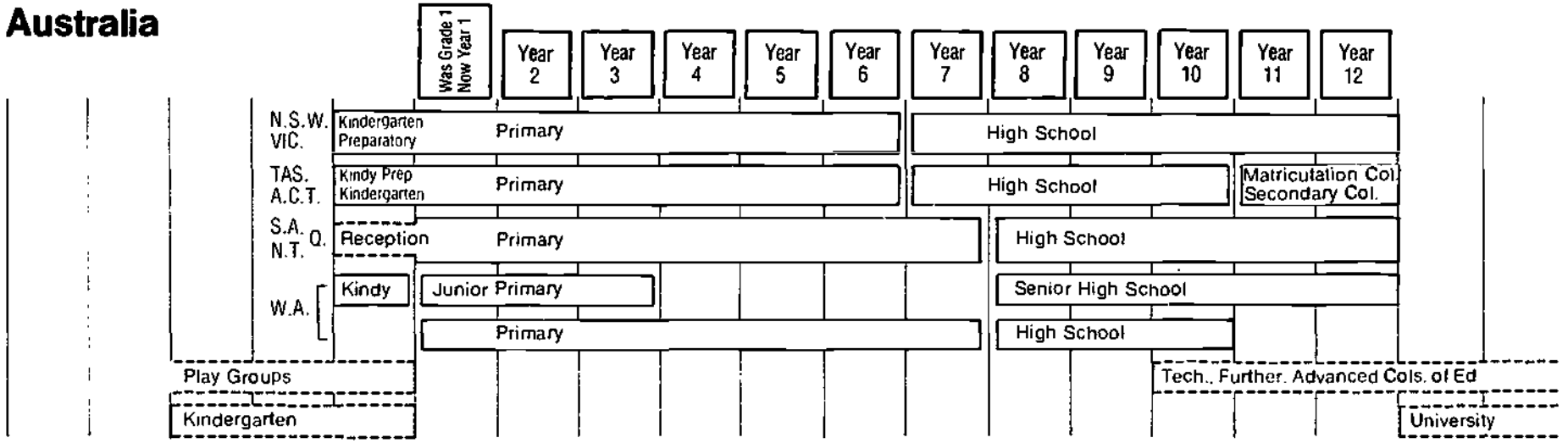
Ages, Stages, Classes, Years, Grades, and Systems compared

Note: This chart gives a rough guide only. All systems are changing and in all cases the chronological ages are approximate.

Approximate Chronological Age

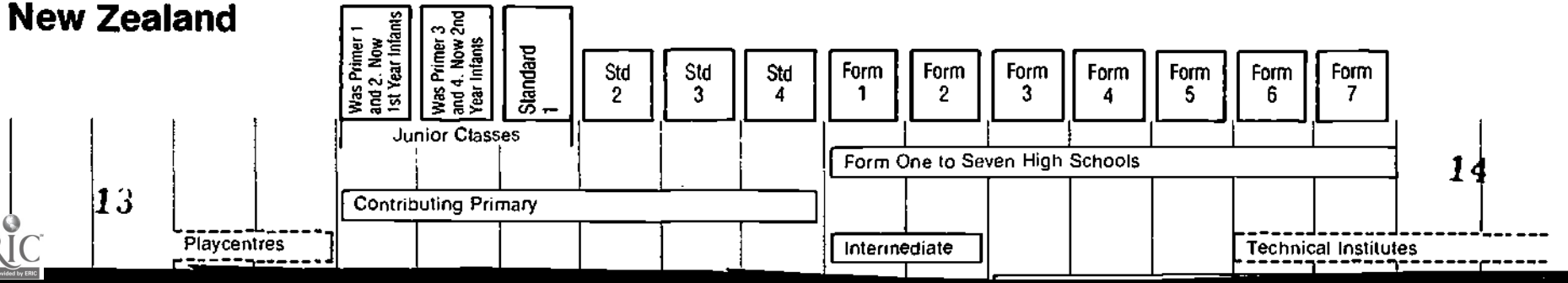
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18

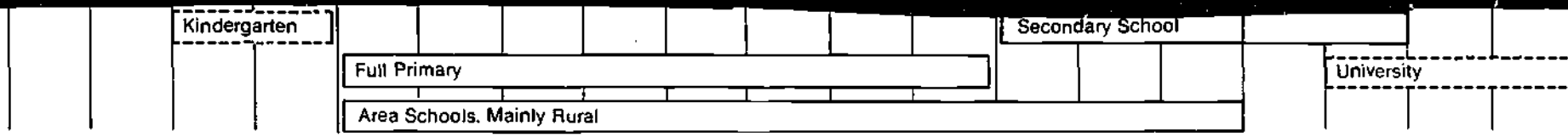
Australia



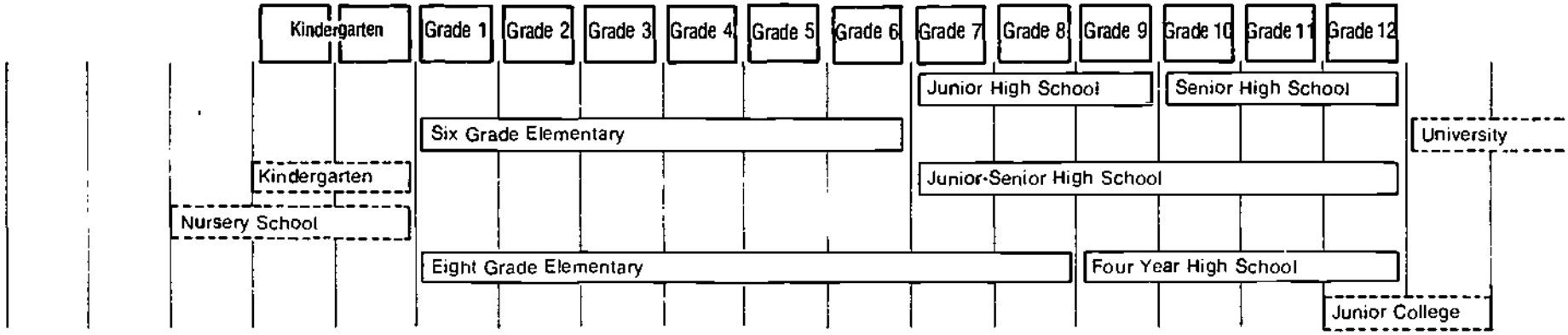
Most States have some State Preschools. Most States have Rural or Area Schools Crossing Usual Primary/Secondary Breaks.

New Zealand



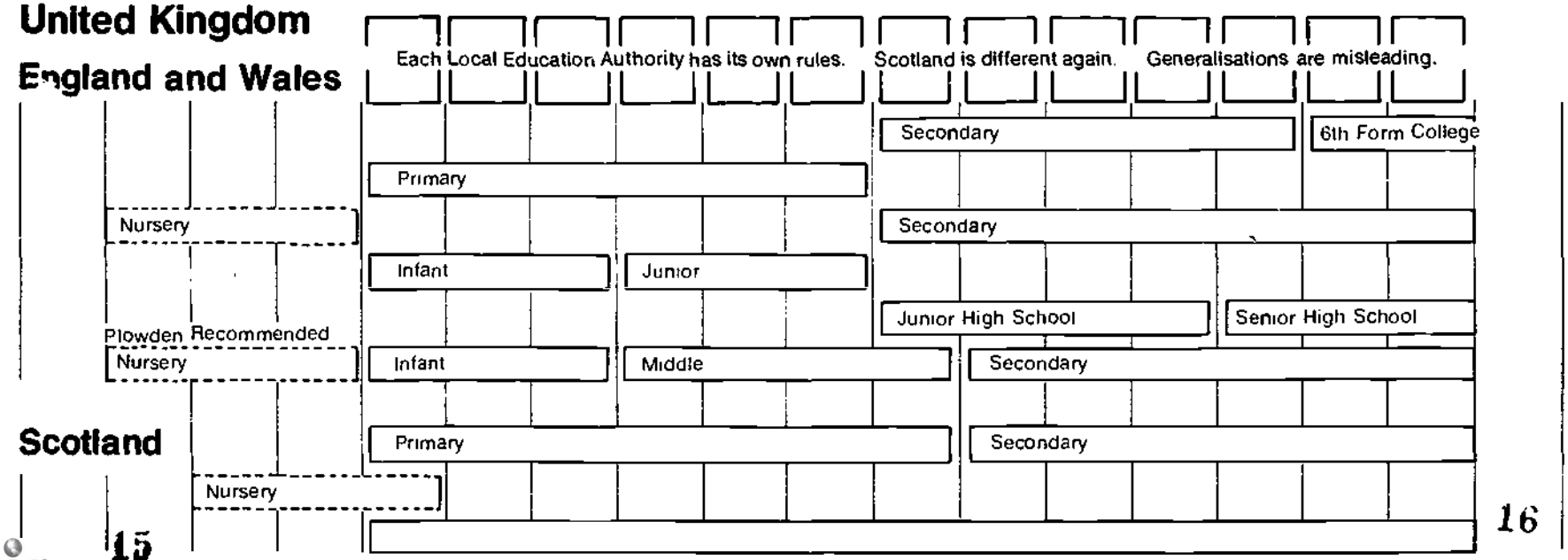


United States and Canada



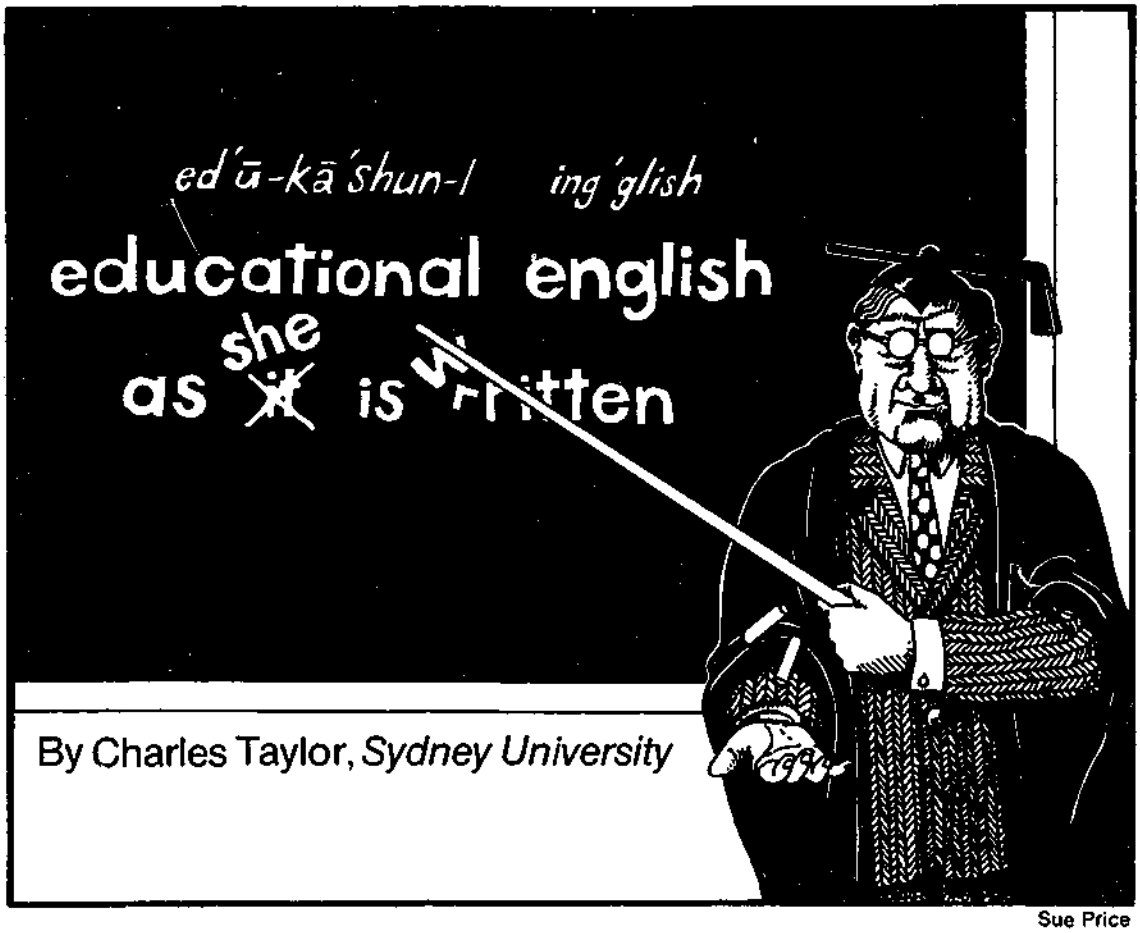
United Kingdom

England and Wales



Scotland

set research information for teachers
number one 1980



Sue Price

What can a teacher, faced by the problem of helping immigrant second-language pupils in his or her class, learn from recent research? One important research project has made us aware of the sort of obstacles which books, as well as classroom talk, are throwing in the path of the immigrant pupil. In this item is described the EMEC (English for Migrant Education Curriculum) research project and conclusions are drawn for everyday teaching practice.

Educational English as She is Written

By Charles Taylor, *Sydney University*

What is the English of school textbooks like? Impressionistic statements such as 'Far too hard', 'There is a shortage of carefully graded texts', 'Not enough solid meat in them these days', are not good enough. This research project, which looks at what is actually read in NSW schools, began in 1975. It paused in 1978 and these are some of the results.

The Methods

All 669 schools in NSW which have courses in years 7 to 10 were sent a one page questionnaire. From subject teachers came information about the most frequently used textbooks. A cut off point of 18 well-used books gave us 700 000 words of running text in six subjects: science, mathematics, history, geography, social studies, and commerce. This was the material analysed in 1975-8.

The analysis was computer assisted and looked at syntax (sentence structure), semantics (meanings) and vocabulary.

The Questions

We set ourselves the task of answering such questions as:

> What is written English like, especially what is it like for non-native readers?

Most obviously it is not like the exercises that are written for foreigners learning English. Those exercises usually consist of single unrelated sentences. Understanding them requires effort for the learner but it is quite a different art from fluent reading. Where a text has been simplified for foreigners it has usually been done by chopping up long sentences into short ones. Yet a string of short sentences is harder to understand than a group of somewhat longer ones written so that the relationships between the parts are made clear by such phrases as 'Nevertheless', 'In the meantime', 'One thing led to another and . . .'.

The problem for the reader is making the right connections. Suppose for example that your eye had caught

'the reader is making'

out of the last sentence. That is a fairly good reading span — bigger than most second-language users in years 7 to 10 could manage. But of what value is it to read those words unless you have the correct divisions between them and the correct links with the rest of the sentence?

The Readers

We found that non-English-speaking learners differ in more than one way from long-resident Australian learners. They tend to respect the textbook more: they tend to give up trying to understand classroom talk; they are shy of spoken interaction in class; and they find that using books gives them more time to understand for themselves without involving them in awkward social problems. For these reasons, and also because more research can be done cheaply with written text, we confined our attention to the possible problems students who have English as their second language would be likely to encounter in their reading.

The Analyses

Our first batch of analyses concerned syntax — the way sentences are held together. We noted that single sentences avoid the problems that can arise when we make a narrative or an argument flow along — their meanings tend to be self contained. Also single sentences are seldom dependent on some situational or psychological context.

> The pen is on the table.
> The car is in the garage.
> The pen is in the box on the table.

These from a typical English-as-a-second-language course. They are not being taught for the content they contain, but for the rules they exemplify — the use of the prepositions 'on' and 'in'. They are specific too. Subject textbooks contain very few such sentences.

> When you have observed a white precipitate forming remove the beaker from the heat using a heat resistant cloth, stir, and filter the mixture using the apparatus you have already set up.

That is a much more common type of sentence, and as you can see it is long, full of connections such as 'when' and it refers to a particular situation in a laboratory.

What is a complex sentence? There is a whole literature of attempts to describe complexity, but the notion that emerged from this study was that it concerns 'interruption'. We used a simple *interruption index* using long interruptions of continuity as the mark of complexity.

'Using a heat resistant cloth'

in the science sentence above is an example of what is, for a second-language learner, a long interruption. Being able to move along a page from one eye-fix to the next and retain the meaning depends on your knowledge of grammar and vocabulary. Vocabulary tells you what the paragraph is about and grammar what it says about it.

We begin by asking what is interrupted. Usually a particular line of argument. It is usually interrupted to take in some supporting evidence or an implication.

> William and Mary, much harrassed by the Monarchists on one side and the Parliamentarians on the other, pursued a middle course.

The next question is how long must the reader hold part of the argument in suspense. Here we need help from the psychologists. They tell us that a wait of

more than 1.1 seconds is a bit too long for proper understanding.

This short-term memory span covers reading up to ten words. That is, if you're an efficient long-resident native reader. But if you're a recent arrival, who doesn't know a lot of English, or if your family doesn't encourage you to read English, you'll miss the boat. That means, the 1.1 seconds has gone before you can catch the drift of the whole sentence. The first bit you read got lost before you finished getting the second bit.

What happens is this. There are invisible relationships between parts of sentences. These are partly structural, i.e., they hold sentences together. But sentences are also tied together even when they are separate.

Look back at the paragraph above, and you'll find a word *these*. The word *these* refers to *invisible relationships*, even though that phrase is in another sentence. We call all these links 'bonding', and it's because good readers get the bonding links right that they know what the sentences mean. Some people study grammar in school still, but the links grammar describes aren't enough. When we read a continuous bit of writing, we need to process all the links. That includes what linguistics call 'cohesion'.

Vocabulary

We looked at an idea called 'collocation'. This vogue word means the closeness of one word to another word in a passage. Take these sentences:

The cat was basking in the sun.

Cardinal Wolsey basked in the warmth of Henry's favour.

You notice that the rather rare word *bask* has to go with some word suggesting heat. So we collect all these words that go together and come up with a cluster of words. These words form a sort of topic pool for vocabulary. You can try this for yourself on some geography book, for example. Notice how each word has other words that seem to be demanded for company. Then collect them into a list, and you've got a vocabulary base for practice material in English.

In our study we used a measure which we called a *span*. Count the words from the keyword — the one you're investigating. Going from *bask* in the sentences about the cat and Henry, you find *sun* and *warmth* are both 3 places away. They are 'at a span of 3'. We used a span of 4 for our research, and you can read some of the results in the Report.

The study also covered word frequency in different subjects and it tried to identify words which do a lot of work. For example, a question like, 'Who won?' said after a soccer match can throw people whose English is weak. This innocent looking question really means:

'Which team scored the most goals, and therefore the winner?'

Then we found that innocent words like *cut* and *area* have different meanings in different subjects. In geometry, one line cuts another. But the line isn't now in two pieces. *Cut* really means 'cross' in geometry. The *area* of a triangle is the amount of space in it. An *area* in geography is just a region. So, 'find the area . . .' What does it really mean? All you native users of English reading this think this is child's play. But not so the second language users.

Grammar

We all hate grammar, but we all use it, or else we couldn't get our ideas across. We looked at grammar, and especially at verbs. Verbs are the key to a sentence. Each verb has a different meaning according to the nouns that go with it.

In order to show the funding authority that we were scholars and not just playing around, we called this sort of thing 'transitivity'. As far as teachers go, it's really just putting verbs and nouns together. Take the word *find*:

(a) Stuart found a lake at the foot of the hill.

(b) Find the value of x in the following.

In (a) because *find* goes with the noun 'lake', we know it means 'discover'. In (b), it goes with 'value', so it now means 'look for'. Keeping those meanings in their right place is a hard game for new readers in high schools.

Language for What?

Is there a special language for science, and one for maths and so on? We didn't find so much of a special language for subjects, but we did find that there was a sort of educational language that tries to be objective.

Objectivity can be achieved by using:

(i) the passive voice: *the island was discovered in 1856*

(ii) *-ing* and *-ed* words: *boiling the water; attached to the top*

(iii) the 'always' tense: water *boils* at 100°C

(iv) introductory *it*: *it is interesting that . . .*

(v) introductory *there*: *there are* many *farms* in the area

(vi) forms with *what*: *what is needed is . . .*

(vii) verbs like *may, could, must*: *this may be the cause of . . .; there must be some reason*

Problems

The passive voice in English has two meanings. Look at:

(a) The solution is placed in the beaker.

(b) Balloons are filled with light gases.

In (a), the passive really means:

(i) Put the solution in the beaker

or

(ii) If you put the solution in the beaker, (the experiment will go well).

But in (b) it means:

People usually fill balloons with light gases.

Verbs like *seem* and *think* often come in funny places in English. Thus:

(a) She seemed not to have heard
(b) I don't think that's right.

In (a) it's not *she* who *seems*. What we mean is: It seemed as if she didn't hear. Foreigners find this really queer. In (b) it's not that I don't think. I do think, and I think that isn't right.

Reading spans we've already mentioned. They are a problem because readers get the wrong words together in a passage and so invent new expressions. In this sentence:

What they found of importance was negligible.

Some readers invented an expression 'find of'. But *of* doesn't go with *find*, it goes with *importance*.

Some conjunctions like *so that* can introduce different ideas:

(a) Place the compass so it faces north (in such a way)
(b) Pull the cord so that the parachute unfolds (with this purpose)
(c) He stopped suddenly, so that the baggage fell off (with this result)

One innocent letter we should all watch closely is the letter **s**. Both verbs and nouns use it in English:

(a) The mist clouds our vision.
(b) The clouds moved slowly.
(c) The cloud's in the way.
(d) The cloud's base was flat.

The comma above the line, the apostrophe, helps us a bit, but even it means two different things here, and we still need to think so as to sort out the meanings.

Simplification

We tried to measure what was simple. We used the idea of *lexicality* and *lexical density* (see the notes) and linguists' ideas like *clause depth* and clause *interruption* mentioned earlier.

We noted that simple words were often more vague. So there's a need in all subjects for specialized vocabulary. This helps sort out the exact meaning.

Some Questions

Are all high school teachers English teachers? Do subject teachers overlook language problems, or hope the English teachers will sort them out? Should we use students' own language as the basis of textbooks, or does this make it impossible to advance further? Is it really possible to teach the English of geography 'situationally'? (i.e., using a real situation, 'the chalk is in the box on the table' often used by English teachers). Is teaching vocabulary enough?

It seems as if many of these questions are answered in part by bearing in mind how we read. As we read, we have expectations that the sentences will be completed. Reading isn't just a passive occupation. If our expectations are disappointed, we get confused. Therefore we must find out exactly what we ought to expect in a school textbook. And that's what the EMEC project tried to show.

An Example

Take the sentence:

> The critical solution state of the ore is related directly to the composition of its crystalline elements.

What is it *about*? The keyword is *ore*, but that word is embedded in the sentence at a point where some readers wouldn't expect it. In particular, foreign speaking readers might not notice it. It's a small word. And yet it's the right place. Our students need a lot of practice in this sort of English.

Students need practice in learning:

(i) to find the links between words.
(ii) to discover which vocabulary words go together.
(iii) to understand objective language.
(iv) to identify keywords in sentences and in passages.

Notes

Some of these references are rather technical, but they repay study:

About Cohesion:

M.A.K. Halliday and R. Hasan: *Cohesion in English*, Longman, 1976.

About Context, or how much work a word has to do:

B. Bernstein: 'Social class, language and socialisation' in Open University: *Language in Education*, pages 102-110, esp. page 106.

About Lexicality and Lexical Density:

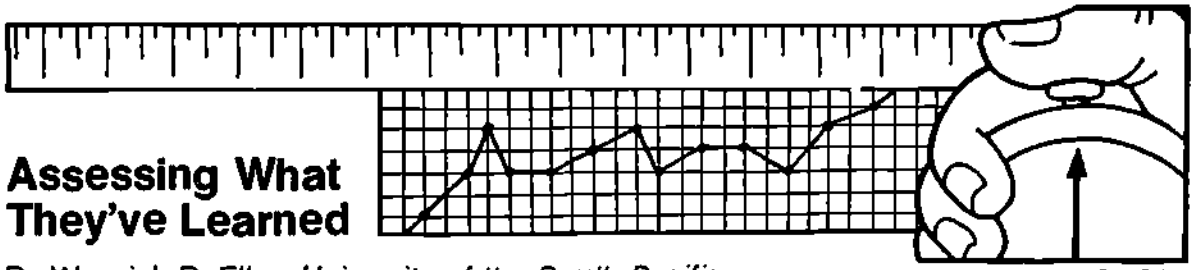
J.N. Ure: 'Lexical density and register differentiation' in G.E. Perren and J.L.M. Trim (eds.): *Application of Linguistics*, Cambridge University Press, 1971.

About the EMEC Project:

C.V. Taylor: *The English of High School Textbooks – Education Research and Development Committee Report No. 18*, Australian Government Printing Service, P.O. Box 84, Canberra, A.C.T., 2600, and obtainable at Government Bookshops.

The Author

Dr C.V. Taylor is Director of the Department of Education's Diploma in the Teaching of English as a Foreign Language at Sydney University.



Assessing What They've Learned

By Warwick B. Elley, *University of the South Pacific*

Sue Price

1. Introduction

Why do we test pupils?

Most teachers spend a lot of time and effort preparing tests, giving tests, marking tests and using tests for one purpose or another. Why? The main functions of tests and examinations in school are:

(a) *Mastery*. Classroom tests are often prepared by teachers to see whether pupils have mastered a particular unit or skill that has just been taught.
e.g., Teacher gives a quick quiz to see whether pupils have learned how to multiply fractions; or know the main events and characters in the book "Animal Farm"; or have learned the main vitamins and the foods that contain them.

(b) *Diagnosis*. Tests are often used by teachers to determine the major weaknesses that a child shows in basic skills.
e.g., In mathematics, to see whether pupils know basic number facts, or can handle zeros, or understand how to divide fractions; in reading, to see whether pupils' difficulty is due to poor vision or hearing, or word attack skills, or vocabulary weakness, or some misconceptions about print.

(c) *Reporting Progress*. Formal examinations are often used to report on the progress made by pupils over a term, or school year. The results are of interest to parents, pupils, potential employers, and teachers in next year's classes.

In addition to these three main uses, formal tests and examinations are often used

(i) to place children in ability groups;
(ii) to select pupils for further education, scholarships, etc.;
(iii) to match pupil materials with pupil abilities;
(iv) to evaluate one's own instruction;
(v) to assist with vocational guidance;
(vi) to determine a child's readiness for learning;
(vii) to undertake research on pupils' abilities, or on teaching methods;
(viii) to determine whether educational standards in a school, or total system are changing.

As teachers, we should be clear *why* we are testing. We should not test just because it is always done. Certainly, tests often do help motivate students to work harder. But the results can be discouraging too, if the results are poor.

In fact, the purpose of the test should affect the kind of test given. Thus, a formal selection examination will normally cover a large number of skills and topics lightly, and most questions should be of middle difficulty level. Mastery tests on a particular unit will be more intensive, with several questions on each of a few topics to see whether they are mastered or not. Some tests should be relatively hard for all pupils, (e.g., Diagnostic tests). Some will be lengthy and formal with elaborate marking schemes (e.g., End-of-year examinations); at other times the teacher will give short, informal quizzes with emphasis on quick feedback. Sometimes the need is for many short questions which can be objectively marked. At other times the teacher's purpose will be better served with a few long answer questions. It is important therefore that we think through our reasons for testing.

2. Qualities of a good examination

Not all examinations are well set. Often they are too hard, or too easy, or not enough time is allowed. Sometimes the questions are vague, or trivial, or provide clues for the 'test-wise' pupil. Many examinations are unbalanced, providing too many questions on some aspects, and too few on others. Such weaknesses means that decisions based on the examination results will be unfair, or pedagogically unsound. How can we tell if our examination is a good one? Two important features are *Reliability* and *Validity*.

(a) *Reliability*

Tests are reliable if they produce *consistent* results, if they produce similar marks on different occasions. If a pupil gains 75% in a reading comprehension test today, and only 50% tomorrow, then the results are not consistent; the tests are not sufficiently reliable to base judgements on. If a pupil is placed first in his class in a test of multiplication and division of decimals today and is 20th in a similar test tomorrow, we can conclude that the tests are not reliable indicators of his ability.

To be reliable a test must normally be *long enough* to minimize the effects of chance factors in the content and skills included in the test. With a short test, a pupil may be lucky, because he happened to know or guess correctly the few questions that were asked, whereas he knew very little about the areas untouched by the test. A standardized test of reading, mathematics or language normally needs at least 40 good objectively marked questions, to reach a satisfactory level of reliability. To make decisions about individual pupils, for placement, or grouping, or diagnosis, a teacher-made test will probably require more questions than this; for judgements about the performance of whole groups, a teacher can get by with fewer.

Just how long a particular test should be, depends on the type of material tested, the amount of supplementary information available, and the importance of the decisions being made. Thus a test of a highly specific skill, such as arithmetical addition, or spelling, or typing, may produce reliable results within ten minutes. If however, we wish to examine a pupil's grasp of a variety of mathematical relationships, or his understanding of a period of history, and then to make decisions about future schooling on the basis of the results, we may wish to extend the test over two or three hours to gain maximum reliability. For such general skills as essay-writing ability, or oral expression, it is commonly found that pupils vary so much in their performance from day to day and from topic to topic, that the only sure way to gain adequate reliability is to test the

pupils on several occasions and several topics, and combine the marks given by two or three independent markers. This may not always be practicable, but we should realize when our results are likely to be fallible.

Other requirements of a reliable test are clear, precise directions, and reasonable time limits. If students are rushed, their performance may not be typical. The questions should be clear and unambiguous, neither too easy nor too difficult; they should normally discriminate well between good and poor pupils, and they should be capable of reasonably objective marking. Otherwise the results will vary according to the values and whims of the marker. If a choice of questions is allowed reliability usually drops, because markers cannot compare answers so consistently.

These are some of the more important factors in determining how reliable a test will be. It is possible to assess the reliability of a test statistically, but that is a topic for another time.

(b) *Validity*

A good test must be valid. This means that, in addition to measuring a pupil's achievement consistently, it should be *relevant to the main objectives* of the course. It should cover the unit or course adequately, sampling each content area and skill in appropriate proportions. If a teacher knows precisely what his objectives are, he can usually tell, by analysing the questions of a test, whether they conform closely to the objectives he has adopted i.e., whether the test has 'content validity' for his purposes.

To illustrate, a 60-item test of addition in arithmetic may be highly reliable, and yet be quite invalid for measuring achievement in a course of modern mathematics which emphasizes concepts, relationships and reasoning. The objectives of the test do not match the teaching objectives. A 3-hour written examination in manual arts may give reliable results. But if it does not require students to show the actual skills they have learned, it will have poor validity. The students who do well on a written examination may not be those who do well in the practical skills.

Again a test of geography which focusses on factual details about populations, areas, climate, exports, capital cities, and the like would produce irrelevant results for a teacher who stressed broad concepts, generalized skills and underlying relationships. A valid test of such objectives may require novel or fictitious situations on which to base questions so that a pupil can demonstrate that he has attained these objectives, regardless of the particular factual details he has acquired.

Sometimes tests lack validity because of 'cultural bias'. Questions may be unfair because they assume that the pupil has had particular experiences which he has not had, or read books which were not accessible to him, or seen films or T.V. programmes which he has not seen. Sometimes the test may be invalid because the print is illegible, or the diagrams unclear, or the paper inadequately proof-read. Such problems may distract pupils and cause changes in their rank order. Likewise, cheating will result in invalid results. If pupils can copy one another's answers, or gain prior knowledge of the questions the results will not reflect their true grasp of the subject assessed.

To ensure maximum validity for his tests, then, it is important for a teacher to spell out, as clearly as possible, precisely what his objectives are, and to build his questions around these, in the appropriate proportions. Tests which develop without such planning often degenerate into factual quizzes of the low-level, isolated, easily testable fragments of the course.

Guidelines for Checking Reliability

1. Is the test long enough?
2. Are the questions clear?
3. Are the time limits realistic?
4. Are the questions of appropriate difficulty?
5. Is the marking effective?
6. Are the instructions clear?
7. Has the choice of questions been kept to a minimum?

Guidelines for Checking Validity

1. Are the questions relevant, important?
2. Have all topics been assessed in appropriate proportions?
3. Have all skills been assessed in appropriate proportions?
4. Are there clues to the right answers?
5. Is the typing and presentation adequate?
6. Have all students had an adequate opportunity to learn the material tested?
7. Is security adequate to avoid cheating?

3. Planning the test

If a test is not well balanced it will not be valid. Therefore, to ensure proper balance, it is a good idea to draw up a plan or blueprint. List the main topics to be covered on one axis and the major skills to be developed on the other.

Thus, a blueprint for a unit on Mathematics might look like this:

	Skills			
Topics	Memory	Computation	Application	Total
Sets	3	—	5	8
Fractions	1	3	5	9
Measurement	2	3	5	10
Decimals	1	4	5	10
Statistics	3	5	5	13
Total	10	15	25	50

The teacher who prepared this plan has clearly decided that the most important objectives in his course are those concerned with applying the skills learnt in new situations, rather than memory work or routine computational skills. Therefore, 25 of the 50 questions are devoted to application. Likewise, statistics is given more weight than the other topics, although all receive some weight.

Planning of this kind should be undertaken in every subject and the weights given should reflect the amount of emphasis given to the topics and skills during the teaching of the unit or course. For example, in Social Studies, the topics to be weighted might be:

Location, Climate, Discovery, Early Settlement, Industry, Transport, Culture. The skills assessed might be Recall, Comprehension, Application, and Evaluation.

An English test might have as its main topics: Written Language, Oral Language, Grammar, Fiction, Poetry. The skills to be tested might be Knowledge, Comprehension, Application, Synthesis (production of original work), and Evaluation.

4. Test questions

Question writing is an art that depends on clear understanding of the subject and of the pupils being assessed, as well as a grasp of the general principles of item writing. It helps, also, to have plenty of time, some imagination, access to other people's questions as models, and an opportunity to have your questions edited by colleagues.

Several kinds of questions can be used, and none is ideal for all circumstances. Written test questions can be simply divided into two types:

(a) *Objective Questions*: These have right or wrong

answers and markers should agree on which are right and which are wrong.
(b) *Subjective Questions:* Essay-type tests in which the pupils must respond to open-ended questions by composing their own answers. There are varying degrees of completeness and correctness.

OBJECTIVE TEST QUESTIONS

(a) *Multiple-choice*
These consist of a *stem*, stating the question, and 4 or 5 possible *options* to choose from.

e.g. (i) What is the area of a rectangle which is 5 cm long and 3 cm wide?
A. 8 sq cm.
*B. 15 sq cm.
C. 16 sq cm.
D. 30 sq. cm.

Note that the "distractors" (A, C and D) should be plausible answers for the pupils who might be unsure.

(ii) If bread is placed in a refrigerator, it will not become mouldy so quickly, because:
*A. cooling slows down the growth of fungi
B. darkness retards the growth of mould
C. cooling prevents the bread from drying out
D. mould requires both heat and light for growth

This question requires the pupil to apply his knowledge of the relationship between temperature and the growth of moulds.

(b) *Matching Questions:*
These consist of two columns of items selected so that pupils can match the words or symbols in one column with the appropriate word or phrase in the other. Matching questions are useful for testing homogeneous sets of facts e.g., matching books with their authors, chemicals with their formulas, words with the parts of speech they represent, etc.

Country	*Capital City*
1. Fiji ( )	A. Rarotonga
2. Tonga ( )	B. Suva
3. Western Samoa ( )	C. Vila
4. Cook Islands ( )	D. Nuku'alofa
	E. Honiara
	F. Apia

Note that both lists should be *homogeneous*, and one should be *longer* than the other. The main fault in preparing these questions is that each list often contains terms which are heterogeneous. They should all be authors, or cities, or chemicals, or parts of speech, etc. If cities are mixed with minerals, and people, and organizations, they provide obvious clues to help the uninformed pupil.

(c) *True-False Questions:*
These consist of a single statement which the pupils are to mark *true or false*, or right or wrong. They are useful questions for a quick quiz, but guessing can be a serious problem with this type of question. This can be reduced somewhat by asking pupils to correct the false statements. e.g.,

		Ring T or F	If F write the correct answer
(i)	25% of 44 is 4	T or F	_ _ _ _ _ _ _
(ii)	The volume of a mass of gas tends to increase as its temperature increases	T or F	_ _ _ _ _ _ _
(iii)	Fiji's chief export is copra	T or F	_ _ _ _

(d) *Completion Questions:*
These consist of a question or sentence containing a blank, for which the pupils must supply the appropriate word, symbol or phrase. These questions are actually "semi-objective", because there is often more than one acceptable answer.

e.g., (i) What is the name of the instrument used to measure temperature? _____

(ii) The device used to tell whether an electric charge is positive or negative is: _____

(iii) What is 25% of 44? _____

Which Type of Objective Test Question Should You Use?

There is *no one best* type of item. All are appropriate at one time or another, but multiple-choice questions are more widely used than others in standardized tests and important examinations. The following advantages are often claimed for multiple-choice questions.

(i) They are more objective and reliable than essay tests or completion questions.
(ii) They make possible the testing of a larger sample of the pupils' knowledge and ability in a short time than does the essay test.
(iii) They enable the teacher to measure process skills as well as recall of simple knowledge. By contrast, true-false, matching and completion questions are largely restricted to simple recall.
(iv) They are easy to mark in large numbers.
(v) They make it impossible for a pupil to gain a high score by guessing.
(vi) Common weaknesses in pupil knowledge and skills can be readily diagnosed by the teacher.
(vii) The questions themselves can be readily evaluated and improved by means of item analysis.

On the other hand, multiple-choice questions do have these disadvantages:

(i) They cannot measure pupils' creative skills, or ability to organize material in a coherent manner. This is particularly important in language, literature, and other expressive subjects.
(ii) They take much time and skill to construct. Poorly prepared questions may produce more invalid results than completion or essay questions.

5. Suggestions for preparing test questions

1. *Essay Questions:*

(i) Specify clearly what is to be included in the answer.
Compare: (Poor) : Write an essay on the French Revolution.
(Better): In not more than 500 words,
(a) Outline the main causes of the French Revolution,
and
(b) Explain why reform could not be obtained without violence.

(ii) Use several short questions rather than one long one.
(iii) Avoid optional questions where possible, as they make marking more difficult.
(iv) Before the test, prepare a model answer, outlining the main criteria and weights to be attached to each.
(v) Mark one question for all pupils before beginning the next.
(vi) Mark without knowing the pupils' names, where possible.
(vii) Obtain independent assessments, wherever you can. The average of two markers is more reliable than the results from one.

2. *Objective Questions:*

(A) GENERAL

(i) Keep your questions brief, simple in expression, and free from complex verbal instructions, double negatives, etc.
(ii) Test only the important facts and skills. Avoid trivial questions, "catch" questions, and irrelevant material.

(B) MULTIPLE-CHOICE QUESTIONS

(i) The problem should be clearly stated in the stem of the question.
e.g., (poor)
Bats
A. drive off harmful birds
B. are enemies of man
*C. eat insects
D. eat rats
The pupils must read all options before they understand what the problem is.

(ii) Use only plausible distractors
e.g., (poor)
The Prime minister of Fiji is
A. Mr Muldoon
B. Mr Fraser
*C. Sir Kamisese Mara
D. The Shah of Iran
Many pupils could guess the answer with very limited knowledge. Names of other prominent Fijians would provide better distractors.

(iii) Ensure that there is only one correct answer
e.g., (poor)
The population of Hamilton is
A. less than 50 000
B. between 50 000 and 70 000
C. over 70 000
D. over 80 000
Both C and D are correct.

(iv) Avoid the stereotyped language of textbooks in the correct answer
e.g., (poor)
The Renaissance in Europe was characterised by
A. a decline in trade
B. many religious wars
*C. an unusual efflorescence of creative talent
D. the loss of colonies

(v) Beware of grammatical clues and verbal associations
e.g., (poor)
The French scientist who discovered the basis for pasteurising milk was
*A. Louis Pasteur
B. Isaac Newton
C. Francis Bacon
D. Alexander Graham Bell
There are two clues to the right answer here. The question should be rephrased.

(vi) Make the correct option the same length as the distractors
e.g., (poor)
Sweets are not recommended for eating between meals as they
A. cause diabetes
B. supply excess energy
C. stimulate the bile
*D. dull the appetite for foods rich in other necessary elements
D sounds right because it is longer, and so makes for a fuller statement.

(C) COMPLETION QUESTIONS

(i) Use a single blank in each question
e.g., (poor)
The "_____ of _____" was written by _____.
A pupil may know the facts required, but be confused by the question.

(ii) Place the blanks near the end of the sentence
e.g., (poor)
_____ is the name usually given to the breakdown of the soil by various processes.
e.g., (better)
The breakdown of the soil by various processes is usually called _____.

(iii) Make all blanks the same length
e.g., (poor)
Villa is the capital city of the ___ _____.
A pupil who was not sure whether to choose Solomon Islands or New Hebrides would have an obvious clue here.

(iv) Make sure that there are a finite number of acceptable answers
e.g., (poor)
Columbus discovered America in _____.
e.g., (better)
In which year did Columbus discover America?
_____.

6. Conclusions

Much more could be said about writing sound questions. However, a careful reading of the principles outlined above and some meticulous editing by your colleagues should make for better reliability and validity than that in a test which grows 'Topsy-like' without planning and forethought.

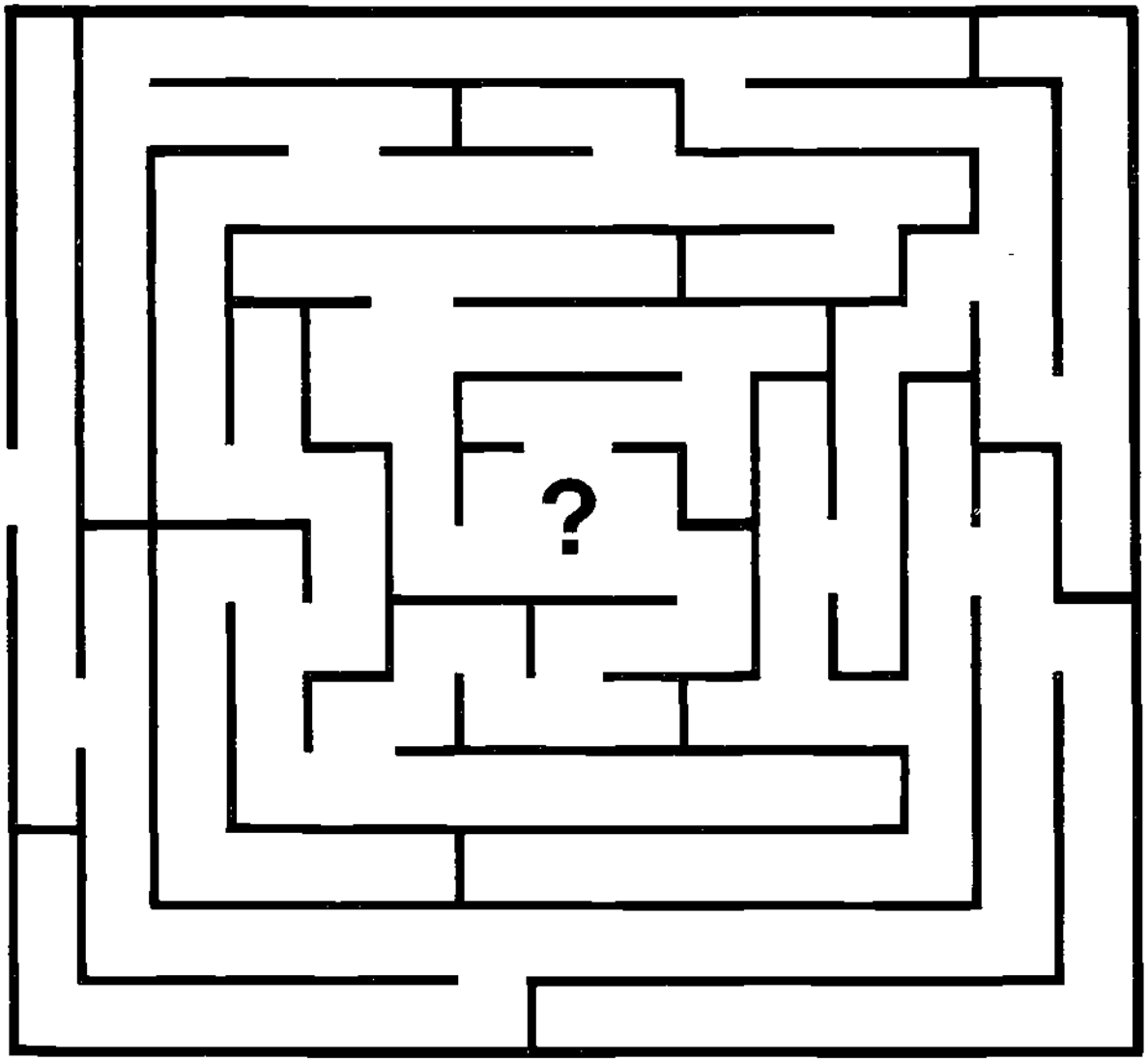
Teachers who wish to improve their assessment skills further can learn much from studying examples of well prepared examinations and standardised tests, and from analysing the results of their own tests, using item analysis. This and other topics can be followed up in such books as:

Ebel, Robert L. *Essentials of Educational Measurement*. New Jersey, Prentice-Hall, 1972 (the test developer's bible).

Peddie, Bill, and Graham White *Testing in Practice*. Auckland, Heinemann Education, 1978. (short, pithy, a practicing teacher's guide).

Queensland Department of Education. *School Assessment Procedures* Titles: 1 An Introduction, 2 The Multiple Choice Item, 3 Assessment in English, 4 Moderation Within Schools, 5 Assessment in English, 6 Assessment in Foreign Languages, 7 Planning a Summative Assessment Programme. 1971–5. Available from ACER.

Izard, J.F. *Construction and Analysis of Classroom Tests*. Melbourne, ACER, 1977.



set research information for teachers
number one 1980

Whatever Happened to Programmed Learning?

By Graham Wagner, *NZCER*

In the late 1950s programmed learning was set to revolutionise teaching. B.F. Skinner started the movement in 1954 with a provocative paper entitled 'The Science of Learning and the Art of Teaching'.

> We are on the threshold of an exciting and revolutionary period, in which the scientific study of man will be put to work in man's best interests. Education must play its part. It must accept the fact that a sweeping revision of educational practices is possible and inevitable.

The scientific approach Skinner had used in his own experimental work with animals could, he believed, be applied to human learning by the use of teaching machines:

> If the teacher is to take advantage of recent advances in the study of learning, she must have the help of mechanical devices.

Six years later Stolurow wrote:

> The teaching machine seems to be a major breakthrough in education comparable to the book, radio and television.

Almost immediately, however, there was a change of emphasis away from machines (hardware) to the programs written for machines (software). The programs were soon (for convenience) presented in textbook form and the machines became less dominant. This was a move in the right direction but even so some reputable scholars still seemed to be mesmerised by the revolutionary possibilities of Programmed Learning and as late as 1962 we were still reading such statements as:

> Programmed Instruction is essentially a revolutionary device in that it has the potential to free man from bondage.

It was easy to see the truth of the arguments Skinner put forward in his attack on the shortcomings in the education system of the USA: other Western countries like Australia and New Zealand were also vulnerable. The education systems were static, tradition bound, content dominated, placing the teacher and the curriculum first and the learner a very poor second. In general, the teaching methods revolved around the teacher's knowledge of the syllabus and previous exam papers. The most common teaching aids were the textbook and a chalk board, and the most recent innovation, in New Zealand at least, was to change the boards from black to green. The syllabuses gave little help to teachers: they provided lists of what was to be taught, but little, if any, help with how one idea should follow on another. Further, in the examination dominated climate, any reinforcement (such as information given by the teacher to the pupils about whether they were getting their work right or wrong) usually came too late, too infrequently, and too unsystematically.

The teacher was in a poor position to remedy any of these shortcomings Skinner suggested without some form of help. However, the new technology available to the teacher, he believed, made it possible to research the complexities of human behaviour in a controlled 'laboratory-like' environment with the help of machines:

> Now, the human organism is, if anything, more sensitive to precise contingencies than the organisms we have studied. We have every reason to expect, therefore, that the most effective control of human learning will require instrumental aid.

Skinner saw teaching machines as a part of an improvement to teaching techniques. In this respect he was saying nothing really new. A machine had been used in history as far back as the middle ages for the training and amusement of knights. Called a 'quintain' it was essentially a figure of a swordsman mounted on a pivot. In one hand the figure had a sword and in the other a shield which served as a target. If the horseman with his lance did not strike the the shield squarely in the centre the quintain would pivot quickly and slap the knight with his sword as he sped by. This form of 'negative reinforcement' was a strong inducement to improve the knight's performance.

Sidney Pressey is regarded as the inventor of the first modern teaching machine although Thorndike is credited with having thought out a way to program textbooks using a machine-like system in 1912.

Researchers investigating the effectiveness of machines and programs have turned away from talk of programmed *learning*, which directs attention at what learners do with their machine or programmed book, and are now looking more closely at what program writers ask learners to do. This emphasis leads them to prefer the term programmed *instruction*. As Hartley, a prominent writer on the movement points out, they are both sides of the same coin.

The Techniques

In the early days of programmed learning the list of techniques and principles advocated was something like this:

(a) The learner must work on his own, with a machine or programmed text.
(b) He must work at his own pace.
(c) He must work through small steps.
(d) These steps must be carefully sequenced.
(e) The learner must respond in some way to each step — usually by writing his answer down, or by pressing an appropriate button.
(f) The learner must receive immediate knowledge of results concerning the correctness or appropriateness of his answer.

An additional principle was also borne in mind, but this did not affect the learner's way of working. This was:

(g) The program must be tried out, and revised according to the results of the trials so that the final version of the program teaches specified objectives to a specified group of learners.

These techniques were simply and practically translated into the following programming steps which writers employed when constructing a program for either a book or a machine:

(1) Performance objectives were clearly defined.
(2) A task (or topic) analysis was carried out.
(3) A teaching sequence was selected.
(4) The program was written (in, for example, linear or branching format).
(5) Then the program was evaluated and revised (if necessary), before it was used in a classroom.

Theoretical Matters

There is now enough evidence to support the use of programmed instruction in selected areas of the school curriculum without spending a lot of time justifying its usefulness. But 'mechanised learning' is a sort of bogy in our educational folklore and it will be as well to look at where it comes from and consider it in the light of what we now know.

Skinner backed up his description of how learning takes place in laboratory rats and pigeons with a theory — that of *reinforcement*. A reward is a positive reinforcement of the action that brought the reward; the action (or response) being what the experimenter, or teacher, is trying to get learnt. A punishment is a negative reinforcement for some response which the 'teacher' is trying to avoid. Liberal minded people were pleased to learn that rewards for the 'right' actions get 'better' results than punishments for the 'wrong' actions. (A little awkward was the discovery that a hungry rat finds food quicker if you give him a little shock [punishment?] as well as the food.)

Rewards can be, for humans, just the pleasure of getting the right answer. So, the argument goes, give all learners the right answer right away, everytime, and

learning will proceed apace. Keep the steps small and the positive reinforcement will come more often and learning will be even faster. This sort of theory, which makes no mention of what might be going on in the learner's mind, is often called the 'behaviouristic theory of learning.'

Skinner's theory came under close scrutiny. One criticism centred around whether studies of animal learning tell us anything about how humans learn. Ausubel in 1963 said:

> The methods of learning employed by animals in mazes do not necessarily correspond to methods of learning children use in grappling with verbal matters in classrooms.

Another criticism was that 'laboratory' results had little relevance to 'natural' learning situations since the simplification of behaviour necessary for laboratory study hides the complexity of human learning to the extent that the results become misleading. Today, studies of human learning favour realism over experimental control in research methodology.

Social and intellectual factors in learning are important. Critics have attacked programmed learning because it is a method that isolates the learner and deprives him of the chance of learning from others. This is a fact and must be taken into consideration when a teacher introduces programmed learning. Isolation can be detrimental to some learners if not used in conjunction with other teaching approaches. Past experience, and the effect it has upon what is to be learned, must also be considered. Similarly the need for creativity to be encouraged and acknowledged.

Although Skinner's simplistic description of how learning takes place has been largely discredited, programmed learning survived as one useful technique of teaching. Several myths about it seem to have survived also. One of them is that programmed instruction is incompatible with humanistic approaches to teaching. Many readers may be surprised to hear that humanist theoretician Carl Rogers in his book *Freedom to Learn* writes that there is a clear place for programmed instruction in those instances in which the student encounters 'gaps in his knowledge, tools which he lacks, information which he needs to meet the problem he is confronting. Here the flexibility of programmed instruction is invaluable'. While advocating caution against using programmed instruction as a substitute for thinking he believes 'that it is one of the most powerful tools which psychology has as yet contributed to the field' (p.141), that is, when it is used flexibly.

Impact on Teaching

Successful though programmed learning has been, the debate surrounding the laws of learning and the impersonal machines made school teachers wary of the technique. Twelve years after it all began there was little evidence that programmed materials were being used extensively in American schools. However, theorists were still experimenting with different methods of sequencing materials.

In New Zealand programmed learning made some headway in industry, a few government departments and the armed services. There were attempts, largely unsuccessful, to introduce the techniques into schools through the universities and teachers' colleges. The Department of Education was not interested. There was loud debate over the impersonal nature of machines, the boring sameness of programs, the inroads of a behaviouristic philosophy which made no allowance for creativity or discovery learning, and these factors helped to turn teachers off the innovations. Although not the only reason for teachers' resistance, the bogy of machines replacing teachers in classrooms was thought of as a very real threat, especially at a time of growing unemployment. In New Zealand it seemed prudent for a small country with limited resources to put its money into 'less controversial' resources such as New Maths, PSSC Physics, Open-plan classrooms, and the like. Teachers were too busy to spend time grappling with a new method of teaching about which there were grave doubts.

Even if there had been no controversy, the time required for training teachers in the new techniques, and the resources required to implement such programmes were in most cases insurmountable hurdles. For the hardy few who managed it, there was always a suspicious inspectorate not kindly disposed towards innovations that handed over the learning to the student and left the teacher doing something other than teaching. To be a teacher you had to be *seen* teaching.

Today

The concept of what programmed instruction is has widened over the years: what started out in 1954 as a relatively simple technique of step by step self-instruction has become a complex systematic approach to teaching in general. The four inter-related steps (1. stating objectives, 2. analysing tasks, 3. selecting a teaching method, 4. evaluating the effectiveness of the teaching and revising to improve) became a flexible curriculum plan enabling teachers to use large scale teaching methods in a way that is now called a *systems approach* to teaching. Those acquainted with computer jargon will recognise 'system' as the word used to describe the way computer experts go about their planning. Indeed the exciting developments in Computer Assisted Learning are an outcome of the programmed learning movement. Examples of the large scale learning packages which can be designed with the help of a 'systems approach' are various: they would include resource and curriculum packages designed to last a term, or even a year, and containing a variety of audio-visual aids and materials, all to be used with a variety of teaching methods. Some 'systems approaches' resulted in large scale curriculum plans, for example, the Keller Plan, the Dalton Plan, performance contracting, the integrated day, 'Sesame Street', 'Man: A Course of Study', to name just a few.

Table 1, on the back page, is a good summary of the changing nature of programmed learning. Researchers started off by looking at the techniques of programmed instruction but now concern themselves with the whole field of educational technology, course planning and curriculum development.

Table 1 The changing nature of programmed instruction
(each cell indicates areas of research)

1954	1961 behavioural objectives task analysis	1968 task analysis behavioural objectives subject analysis	1974 systematic analyses and constraints levels of objectives techniques of structuring material
small steps overt responses reinforcement self-pacing	small steps logical sequencing active responding immediate feedback self-pacing	task structures and related teaching strategies presentation as a communication problem	task structures and related teaching methods, strategies and tactics presentation as a communication problem; problems of learning from text, diagrams, etc.; appropriate instructional decisions individual differences and methods of learning
revision	evaluation (i) revision (ii) studies	evaluation : revision	evaluation of (i) objectives and methods (ii) techniques: retrospective and concurrent decision making evaluation as illumination
			installation and implementation inservice teacher training

The Contribution of Programmed Learning to Education

With the blurring of the boundaries between media, it became necessary to look beyond the teaching programme and think instead of programmed teaching. This is where 'systems' technology took over from the machines and this change of emphasis may prove to have been programmed learning's greatest gift to education.

However, we got more than a re-emphasis on objectives and evaluation from the experimental work of programmed learning enthusiasts — we got an introductory understanding of educational technology which will help us cope with further technological advances and the complexity of learning; most computer assisted instructional (CAI) programs are, after all, just very sophisticated programmed learning sequences. We are much wiser now about the limitations of programmed instruction. We are wiser too about theories of learning. We accept that programmed learning is only one teaching method out of many and that it works best when the teacher is part of the programme. But we also know that the skills we gained in producing programs are transferable to the whole field of education under the heading of a systems approach to education.

The education system of the future will look very different from the present one: it will be different in large part because of Skinner and the programmed learning movement.

Notes

Skinner's provocative paper, 'The Science of Learning and the Art of Teaching' is found in:

Lumsdaine, A.A. and Glaser, R. (Eds.) *Teaching Machines and Programmed Learning*, Washington, N.E.A., 1960.
The quotation by Stolurow is found in the same book.

The quotation 'Programmed instruction is essentially a revolutionary device...' is from:

Schramm, Wilkin, (Ed.) *Four Case Studies of Programmed Instruction*. New York, Fund for the Advancement of Education, 1964.
The second quotation from Skinner is from the same paper as is mentioned above.

For criticisms of Skinner's ideas see:

Annett, J. 'Psychological Bases of Educational Technology' in Budgett, R. and Leedham, J. (Eds.) *Aspects of Educational Technology VIII*, London, Pitman, 1973.
Hartley, who is quoted for his good summaries of techniques and the changing nature of programmed instruction — see the section **Techniques**, and Table 1 — can be read in:

Hartley, J. 'Programmed Instruction 1954–1974: A Review', in *Programmed Learning and Educational Technology*, Vol. XI., pp.278–91.

Item 11

set
number one 1980

In both Australia and New Zealand there has been a great deal of interest in the effects that external examinations have on the subjects taught and the way they are taught in secondary schools. In 1970 Queensland took the plunge, scrapped a final internal exam and introduced internal assessment. How has the experiment fared?

Assessing Internal Assessment

By Ann Lamont and Bruce McBride,
Department of Education, Queensland

Until 1972, entrance to tertiary institutions in Queensland and selection into many occupations was determined by the applicant's results in an external examination. This examination, which was administered by the University of Queensland, dominated the final years of secondary education in Queensland.

After much debate and the rejection of the 'Bassett' Committee's recommendations in 1969, the Radford Committee was set up to review the system of examinations. In 1970, the gordian knot was cut by the Queensland Parliament: internal assessment was introduced progressively starting at the Year 10 level in 1971 and reaching the Year 12 level by 1973.

In 1974, a major evaluation, the "Schools Under Radford" (SUR) study was made, and in 1978 the 'MAPS' study was undertaken.

This paper describes the major aims and findings of the MAPS Study conducted by the authors for the Department of Education, Queensland.

Aims

The major aims of the study were:

- to map the decisions that have been made about assessment procedures;
- to find out how satisfied those who actually do the assessing are;
- to compare the assessment situation in 1978 with the situation as it was in 1974.

Semester 4 at the Year 12 level was chosen and three subject areas (English, Mathematics and Foreign Languages) in a representative sample of 12 State secondary schools were investigated.

Data were collected by the researchers during interviews with school inspectors, principals, subject area co-ordinators and Year 12 teachers. Also, in each school, group discussions between the researchers and a class of Year 12 students provided data about students.

Background

The third aim of the study was to compare the situations in schools in 1974 and 1978. Data had been collected in the 'Schools Under Radford' (SUR) study conducted jointly by the Department of Education and the Board of Secondary School Studies. This study had not investigated assessment procedures in detail but was concerned with a global evaluation of the secondary school system which operated after the implementation of internal assessment.

The subjects investigated were chosen to represent a wide range of characteristics. English and Maths I were taken by large numbers of students. The skills developed in these subjects differ from each other and are used by students in a wide range of other subjects. Foreign Languages were chosen because each foreign language is studied by a small group which, it is purported, is drawn from a relatively elite group of high achievers.

The final year of secondary schooling, Year 12, was chosen because:

- the results of three different measures of achievement are given to students in Year 12 (moderate school assessments, the Tertiary Entrance score and school reports); and
- achievement measures provided to students in Year 12, whether internal or external, are likely to have a more significant effect on students' future career choices than measures provided in earlier years.

Moderated School Assessment

For each of the four semesters in years 11 and 12, the last two years of secondary school, students are given a rating on a seven point scale for each subject studied. An elaborate system of moderation has been set up by the Board of Secondary School Studies to ensure that ratings given by different schools are of equal worth.

The moderation system involves meetings of the teachers involved in each subject. This system is described in detail in information kits distributed by the Board.

The Tertiary Entrance Score (TE Score)

The TE Score is based on the Special Subject Assessment (SSA) and the Australian Scholastic Aptitude Test (ASAT). An SSA is an achievement measure given by the school to each student in Year 12 for each subject studied in either Year 11 or Year 12. The calculation of the TE Score is described in detail by McGaw, Warry and McBryde (1976). The main steps are:

1. The SSAs for each subject are scaled to the mean for the students taking that subject in their school. This compensates for the differences in ability levels of students taking different subjects.
2. The aggregate of the scaled scores for the highest 20 semester units is calculated.
3. The aggregate scores for students in each school are rescaled to the ASAT mean for the school. This adjusts for different ability levels between schools.
4. The rescaled aggregate scores of all schools are used to rank all students in Queensland in order of merit. TE Scores are then assigned to students in accordance with their position in the order of merit list.

School Reports

The type of information appearing on school reports and the frequency with which they are issued is determined by each individual school.

Results

Summative Assessment

(Assessments made during or at the end of a course of study which contribute to the final grade or rating.)

Information was obtained about the structure of the summative assessment programme. Changes desired were noted. The decision-making processes were also investigated. The major topics studied were:

weighting on various assessment elements; weighting on modes of assessment used; weighting on types of item used; use of common tests (within and between schools).

Weighting

The weighting given to various assessment elements, modes of assessment and types of item varied considerably between subjects and between schools. However, teachers were largely responsible for the choice of what was used, and consequently were generally satisfied.

In Maths I, ratings were based almost entirely on formal tests under examination conditions, whereas in English and foreign languages the ratings were based on a large number of assessments of a widely varying nature.

Use of Common Tests

Within all schools, common Maths I tests were used where two or more classes were studying the same unit. Many mathematics teachers indicated that common tests between schools would be desirable so that schools' results could be compared.

Common assessment for all Year 12 English classes within a school was not universal and the amount of common testing varied markedly between schools. Most schools, however, used more than 70% common assessment between classes. Almost all English teachers indicated that they did not want common testing between schools.

No school in the sample had more than one class studying the same foreign language at Year 12 level. Some foreign language teachers suggested common tests between schools for parts of the course, such as grammar and comprehension.

There was a sharp contrast between the Maths I teachers, many of whom wanted a return to an external senior examination and the English and foreign language teachers, none of whom wanted a return to full external examinations.

Overview of Summative Assessment

In general, teachers were satisfied with the present end-of-the-course (summative) assessment program for their Year 12 classes. Those who did desire change indicated that they wanted a small change in emphasis rather than a change in its structure. In English there was some support for an increased emphasis on oral work and other communication skills with a decreased emphasis on literature.

Most decisions concerning the structure of the summative assessment program were made by the teacher in charge of the subject, usually in consultation with the other teachers. In all cases, the principal could over-rule these decisions, but it appeared that this power was seldom exercised.

Formative Assessment

(Assessments made while a course is in progress to show gaps in knowledge.)

Two of the aspects of formative assessment probed in the study were:

- the use of summative assessment for formative purposes, e.g., middle of term exams used to diagnose what a pupil needs to learn;
- the balance between summative and formative assessment.

The Use of Summative Assessment for Formative Purposes

The major purpose of the summative assessment program is to provide measures of achievement at the end of the pupil's school life, but it can also provide some feedback on student performance. Three aspects were investigated; delays, students' attitudes to marks, and the making of corrections on marked scripts.

Teachers and students generally agreed that the delay before getting Maths I and foreign language papers back was satisfactory. Delays in English were longer and considered unsatisfactory by students in about half the schools. In contrast, almost all English teachers thought the delay was not too long.

Most students indicated that they were more interested in where and why they lost marks than in their overall scores. However, teachers felt that students were more interested in the scores!

In some cases, students were allowed to make corrections on their papers while in other cases the papers were kept 'clean' for moderation meetings. Neither teachers nor students appeared to be concerned with this issue.

The Balance Between Summative and Formative Assesment

The majority of inspectors, principals and teachers indicated that they were satisfied with the present frequency of summative assessments. While students were reasonably satisfied with the summative assessment load in these three subjects, some informal comments by students indicated that those students doing mainly humanities subjects had a heavy work load. (Data relating to these subjects was not sought in this study.)

Special School Assessments (SSAs)

It appears that very few teachers have much knowledge of, or interest in, the procedures used for these calculations in their school. Usually major responsibility for the calculation is given to one or two persons (generally, the principal, deputy principals, head of the mathematics department or a mathematics teacher) with a small degree of participation required from the other subject heads. While principals reported that input from 'other teachers' occurred frequently, these teachers felt that an input from them was seldom sought.

School Reports

The contents of the end of semester report varied markedly among schools. All schools provided at least one achievement measure, and all except one school provided at least one affective measure.

Comparisons between 1974 and 1978

Uses of Formative Assessment

Data collected in 1974 in the 'Schools Under Radford' study indicated that most teachers held very strong

views about the emphasis on summative assessment. They felt it prevented them from devoting any time to formative assessment. This contrasts sharply with the present study which indicates that, in 1978, formative assessment was frequently used, and that over-testing is no longer a serious problem in the subjects investigated.

Flexibility in Assessment Programs

The results of the 'Schools Under Radford' study suggested that teachers in 1974 were not taking advantage of the flexibility in assessment procedures afforded by the implementation of the Radford Report in 1971. By contrast, in 1978 some teachers particularly those of English and foreign languages appear to have capitalised on the flexibility available through internal assessment for modes of assessment and types of item. However, moderation meetings still base their judgements mainly on written work. This emphasis on written work still appears to be constraining teachers to limit their use of oral assessment techniques.

Mechanics of the Moderation System

In 1978, respondents were generally satisfied with the procedural aspects of the moderation system whereas this was not generally true in 1974.

Principals interviewed appear accutely aware of the problems of over-testing and many are taking steps to ensure that over-testing does not occur.

When asked if they would like to change the emphasis on summative and formative assessment, many teachers indicated a preference for a small shift from summative to formative, but generally they did not see the summative assessment program as dominating their teaching. Most teachers in the subjects studied indicated that they were able to integrate formative assessment into their teaching programs. Half of the principals and most of the inspectors wanted to increase the amount of formative assessment.

Validation Procedures

Most principals and teachers indicated that they were satisfied with the quality of the assessment instruments in their school or subject. However, on being questioned further, most of these indicated that their view was based purely on their own perceptions rather than on any specific validation procedure. In contrast to the satisfaction in schools, only one of the six inspectors interviewed was satisfied with the quality of tests used in schools.

Achievement and other Exit Measures

Data was collected on the following aspects of measuring and reporting student achievement: moderated school assessments; special school assessments; and school reports.

Moderated School Assessments

About two-thirds of the respondents were dissatisfied with the moderation system. Many English teachers wanted an alternative which decreased the emphasis placed on moderation, giving schools more responsibility for maintaining standards. However, many mathematics and foreign language teachers desired some form of external assessment.

A considerable number of respondents have doubts about whether schools can be compared accurately and marks adjusted accurately. Many teachers felt that once ratings had been given in Semester 1, it was very difficult to adjust the distribution upward in subsequent semesters if some students showed a marked improvement in their achievement.

Conclusions

The major findings are:

Teachers are generally satisfied that mechanical aspects of assessment and moderation are functioning fairly smoothly.

There is substantial concern that comparability between schools is not being achieved — that the system of moderation is not ensuring that results are fairly scaled to reflect the standards in different schools.

Schools are utilising the flexibility available to them in their choice of modes of assessment.

Respondents generally felt that summative assessment did not dominate their assessment program and that they were generally able to integrate formative assessment into their teaching program.

Only half of the schools used standardisation effectively when combining semester scores during the assigning of Special School Assessments which are sent to the Board for input into the Tertiary Entrance (TE) Score calculations.

Between 1974 and 1978 there have been considerable improvements in: the use of formative assessment; the use of appropriate modes of assessment; and the mechanics of the moderation system.

Overall, there is greater satisfaction with the system at present than there was in the years immediately following the implementation of internal assessment.

Further information concerning the studies may be obtained from the authors c/- Research Branch, Department of Education, Queensland, P.O. Box 33, North Quay, 4001.

Catering for Individual Differences

set

Standard 3, Newtown, 1935

After Play: 11.05 Everyone doing the 7 times table

11.10 Everyone doing the 8 times table

11.15 Everyone doing the same 20 mental

11.20 Everyone doing the sums on page 45, pens down and fold your arms when you have finished.

It was too slow for half the class.
It was too fast for half the class.
It was just right for nobody, well, maybe, for one or two.

It was the class trap. We put the children in classes: but only one or two of them fitted.

We still put children in classes. But we now believe that every child has a right to be taught the things that fit his or her needs, and taught them at the pace and in the manner that fits his or her abilities. These are new responsibilities. The techniques to put them into practice are still being worked out. In this item we look at two very different techniques: **Student Team Learning**, and **Self Paced Instruction.**

Student Team Learning

By Paul Power, *Melbourne State College*

'I can't go on vacation now. My team needs me.' That's what 9-year-old Lisa at Old Forge School, Maryland, said when her parents suggested a trip to Mexico. And the team she was referring to was her learning team of five classmates in her Year 4 (Std 3) class.

Lisa's class is trying out Student Team Learning in three variations, each creating the same kind of peer support, excitement, effort, and camaraderie normally associated with team games. The variations are called TGT, STAD, and JIGSAW. The experiment is under the Centre for Social Organisation of Schools, John Hopkins University, Baltimore. They have spent eight years developing their own techniques and adapting the Jigsaw technique of Elliot Aronson of the University of California, Santa Cruz.

Lisa's class is one of many in the evaluation, which has included thousands of students. Student Team Learning techniques have equalled and in many cases surpassed traditionally structured classes in promoting learning, particularly in basic skills. In addition they have been shown to have positive effects on race-relations in schools — that was a major reason for the development of Jigsaw — and on students' concern for each other, and on self-esteem.

Further, the techniques are easy and inexpensive.

TGT (Teams-Games-Tournaments)

Like the other two techniques, TGT is designed around the learning team which is composed of 4 or 5 boys and girls of different levels of ability and social backgrounds. The teacher assigns the children to teams, each team choosing a name. Teams stay together for at least 6 weeks to build cohesiveness and team spirit.

After the teacher has presented a lesson, worksheets are handed to each team. The team members then study together trying to make sure that each member knows the material. The teacher is free to co-ordinate, assist individual pupils who are having difficulties, supervise the teams and so on. She acts like a head-coach but all the players help their teammates polish their 'game'.

At the end of the week, or the end of a learning unit, the teams clash in a tournament covering the work they have been boning up all week. The games use simple objective-answer questions taken usually from reading and language programmes. 1. *Grass* goes with *ground* as *bark* goes with (a) dog, (b) tree, (c) wolf; 2. *Sleepy* goes with *tired* as *confused* goes with (a) mad, (b) afraid, (c) bewildered; 3. Closed tightly = (a) clipped, (b) clenched, (c) cleaned; 4. The farm belongs to —— family. (a) their, (b) there, (c) they're. The tournaments are well organised with each child competing against others of similar ability from the other teams — winning is equally valuable so low ability students have just as much chance of contributing to their team's success. There are no individual losers since only the team scores count. The winning team is announced in the weekly class newsletter. Questions come in random order, challenging is allowed with penalties if the challenger is wrong.

Consistently high scorers in the tournament are 'bumped' to meet opponents of higher ability next week. Low scorers are moved to meet their peers. Aggressive children take to the system better than others but the system encourages co-operation and interaction across ability levels and across racial and sex lines.

STAD

Student Teams Achievement Divisions uses the same team structure as TGT but instead of the tournament the students take quizzes and short tests individually. If you score as well as usual you get a maximum score for your team, and there are bonus points for those who do better than usual. STAD takes less time and is easier to organise than TGT.

JIGSAW

Each team is given the same four or five pieces of work — usually biographies, short stories, book chapters, and so on and each team member becomes an 'expert' on one piece. Members of different teams who have the same topic meet in 'counterpoint' groups, for about one-third of the time available, to work over the material. Then they return to their original groups and pass on their information to the other members. There is no system of inter-team competition for marks in JIGSAW. Groups should be heterogenous — as mixed as possible — with best friends and worst enemies in separate groups.

There is mixing of abilities, sexes, and ethnic origins, and a great deal of practice in reading for meaning, verbalisation, listening, and explaining skills. Enjoyment of school, co-operation, self-esteem, concern for others and the number of black-white friendships (compared to control classes using ordinary teaching methods) grew substantially.

TGT and STAD are most appropriate for teaching basic material: JIGSAW is more appropriate for subjects such as literature, history, social studies, and the concept-learning aspects of science. Low-cost materials are available which provide sufficient detail for teachers to use the techniques with their own curriculum material (US$3.00). Specific curriculum material has been prepared for language arts, mathematics, reading, social studies, and nutrition for years 3 to 8. These too are inexpensive (US$8.00 per set) but are geared to American schools. Although not directly adaptable in Australia and New Zealand they may provide useful guides to teachers who want to develop their own material for use with the basic techniques, which do offer much promise.

Letting students work on their own: an experiment in 4th Form Science

By John Longbottom, *Aranui High School*

In 1978 I was faced with a particularly 'deactivated' 5th Form science class. To catch even half the pupils' attention was difficult, to hold it was impossible. Several pupils would do no work at all unless someone stood directly behind them. In a desperate attempt to gain more time to circulate and prod along the more reluctant class members I transcribed five lessons, including instructions for experiments, explanations and notes to be copied, onto worksheets. Then, making maximum use of the Hawthorne effect, I told the class it was the subject of an important experiment and presented the sheets to the pupils as a new way of working. The responses ranged from neutral through to very encouraging, with one pair of pupils who had done virtually no work all year suddenly seeking me out to ask questions on the text and experiments.

Later I had the opportunity to try out such an approach under more controlled conditions, and attempted to isolate and measure the effect of allowing students to pace themselves through lessons.

The Programme

In a study, undertaken as Research Affiliate at the University of Canterbury, I prepared worksheets for a five-week block of individualised work in science and tested them under controlled conditions. The results indicate that pupils tend to work harder when left to themselves and their achievement at least matches that of pupils taught by a teacher.

About fifty objectives were developed for Section 4 (Force and Mass) and Section 22 (Forces in Action) of the New Zealand Form 1-V Science syllabus. Based on these objectives, tests and worksheets were written covering a five-week block of work (20 × 1 hr periods). The worksheets were not difficult to write as they contained nothing different in content or approach to that of a normal lesson.

Four middle-ability 4th Form classes worked through the first two-week block and then sat a post-test. Two of the classes (controls) then reverted to 'normal teaching' while two experimental classes continued with the same material in worksheet form. A second post-test was held at the end of five weeks and the results of experimental and control classes were compared (with the results of the first post-test being used to adjust for any initial difference between the classes).

A questionnaire was used to gauge various aspects of pupil attitude towards the method.

The Results

Pupils achieved equally well whether they worked under self-paced or teacher-paced conditions. The questionnaire clearly indicated that pupils put in more work under the self-paced conditions, even when they had a negative attitude towards this method of working.

Commentary

Using the worksheets under more normal conditions after the experiment was over, and where there had been free interaction between pupils, teachers and worksheets, I have encountered a generally positive attitude, with most pupils expressing the desire to use the same method for at least a proportion of their science course. No one method will ever suit everyone and some pupils will dislike working on their own because they miss the security of a teacher telling them what to do or because they have difficulty getting themselves organised. However, it is hoped that, if made to work in this mode for at least some of the time, these pupils will become a bit more self reliant and a bit more organised.

Teachers commented on the improved work habits and attitudes of classes working on the material and particularly on the smooth and rapid start to the lesson as pupils came in, collect the material and begin working. There is an obvious advantage in the method for dealing with pupils who are absent for a time.

I had expected that the pupils working through material by themselves would be much slower than those in classes paced by the teacher. The surprising find was that the teachers reported that it was in their 'ordinary' classes that they had difficulty in covering the work in the scheduled time. This could be due to improved work habits by those using the individualised material — there is increased efficiency when a pupil spends time only on those parts he has difficulty with.

Although the method obviously enables teachers to spend more time with small groups and getting to know individuals, an additional comment made by one teacher was that it was the first time he had been able to stand back and watch the class as a whole working and interacting.

Teachers enjoyed the feeling of being organised ahead of time and also a non-specialist commented that using the programme enabled her to teach fully and enjoy a topic that she normally skipped through quickly.

The material used in this study contained no unusual material or equipment; it is 'ordinary' and as such could be produced for any science topic by any school.

However, teachers may hesitate to embark on the preparation of such materials because of the time required. There are long-term savings once the material has been produced, and the load can be reduced by sharing it amongst the members of a science department and teachers can produce material for topics in which they have paricular strengths.

There is also the problem that the school has to find the cost of providing printed material. We provided each child with about 10 sides of A4 per week (excluding tests) and it seems unlikely that this could be reduced

significantly. One possibility is to extract all those sections on which pupils write and produce this on separate sheets. However, this would provide added complications to the instructions and those sheets which were collected back may well have a very limited life anyway. With the cost of text books rising rapidly many schools have had to move from pupil-texts to class-sets which can be circulated around several classes; the use of printed material fits nicely into such a scheme and some saving on text book expenditure may be made.

A complaint about using prepared material is that it prescribes the teaching programme and leaves no room for teacher innovation, but this can be avoided by using the prepared materials for certain topics only or for the core work in each topic. Providing there is no attempt to build a completely integrated science programme of prepared materials, teachers should have plenty of scope for using their own initiative and ideas and pupils will not become bored by a single method of working.

The Worksheets

The worksheets were designed to follow the course of a lesson and contain:

1. blocks of information;
2. questions that either:
   (a) asked pupils to speculate before reading the next block of information,
   or
   (b) led pupils to a conclusion via a set of cue questions;
3. instructions for performing experiments;
4. summaries of theory or experiments that were to be copied by pupils into their books.

There are a number of problems in producing worksheets:

1. Language difficulty: a conscious effort has to be made to keep simple the general language (as distinct from any necessary technical language).
2. Feedback on progress: a major criticism was that pupils had little information on how they were progressing until the actual tests. A number of built-in, self-administered tests is probably required.
3. Form of the worksheet: there seemed to be a general tendency for the less able to enjoy copying out material from the worksheets into their books whereas the more abte found this rather frustrating and would have preferred a fill-in-the-blanks style of sheet which would let them work faster.

Under the experimental conditions, pupils were asked to restrict their co-operation with neighbours to (a) the gathering of data in experiments where more than one person was required, and to (b) asking small points of clarification of the textual material. These restrictions over five weeks led to a boredom effect: four times as many pupils in the experimental classes expressed a negative attitude towards the self-paced style of teaching. Teachers were similarly asked to restrict their activity to making equipment available, answering small points of clarification and, only when a pupil was obviously wasting a large proportion of his time or distracting others, act to control a pupil's activity. Teachers also expressed some frustration at being restricted.

The control classes worked through the same material under teacher direction with textual material and instructions being read and displayed on an OHP or at times written on the blackboard. Teachers were asked to restrict themselves to the textual material provided and not to encourage questions that might lead to them giving large-scale interpretations or explanations.

Teachers with control classes reported frustration with these conditions so regularly that the extent to which they managed to maintain the conditions for the whole time must be questioned. It is very probable that at least on some occasions they provided explanations and materials in addition to those on the sheet — one is only human, a child's genuine request for information and help is hard to refuse.

Notes:

Two good articles about Student Team Learning are by Judge, Jean, in *American Education* of December 1978 and
Pearson, Craig, in *Learning* of March 1979.

Information on STL (including an overview film strip and cassette tape at US$15.00) is available from Dr Robert Slavin, Student Team Learning Project, Centre for Social Organisation of Schools, The John Hopkins University, 3505 N. Charles Street, Baltimore, Maryland, USA 21218.

Jigsaw is described in several articles, for example, by Aronson, E., in *Psychology Today*, Vol. 8, No. 9, 1975. Further information on the technique is available from Dr E. Aronson, University of California, Kresge College, Santa Cruz, California, USA 95064, particularly in Bender, R.D., Lewis, K.E., Schwartz, J.C. *Talks to Jigsaw Teachers:* A teacher's guide to the implementation of the jigsaw co-operative learning method, 1976.

Paul Power, PhD, is a lecturer in the Department of Educational Psychology, Melbourne State College, 757 Swanston Street, Carlton, Victoria, Australia 3053.

John Longbottom is Head of Science, Aranui High School, Shortland Street, Christchurch 6, New Zealand. He is able to supply examples of the science worksheets he has described.

---

set research information for teachers
number one 1980

The Class Size Issue Rides Again

Starring Gene Glass and Mary Lee Smith, with a cast of thousands — the researchers and the researched.

Our Story Opens to Reveal that:

Teachers are almost universally agreed that large classes "are exhausting, a cause of frustration, and a reason for failure in basic subjects". Their views about what constitutes a "large" class have changed considerably in recent years, however. The same author quotes figures for England to show that the optimum size of class that teachers thought most desirable was regularly about 3 to 5 pupils less than the current average! Parents tend to support the teachers' views about the desirability of small classes and frequently send their children to private schools for just this reason. Both teachers and parents argue that with smaller classes there is less strain on the teacher, and he can give more attention to individual pupils, and so improve achievement levels. Unfortunately, the research shows that the situation is not nearly as simple as this.

The Plot Thickens: A Brief Review of Some Research Studies

The first experimental investigation was carried out by the American, J.M. Rice, in 1902. He studied the arithmetic achievement of 6,000 children in grades 4–8 in relation to size of class, amount of instruction time, and several other factors, and found that class size had no bearing on the results. He repeated the experiment with a test of language in the following year, with the same outcome. Rice concluded that there was no relation between size of class and the results, that some of the best work was done in the largest classes, and some of the poorest in the smallest classes.

Countless similar investigations have been conducted since Rice's day, but many have been poorly designed and executed. Writing in 1954, Blake reviewed 267 published studies on the subject, and found only 22 that met the criteria of "scientific adequacy". Eleven of these had been evaluated for achievement test scores, and of these, 5 favoured small

classes, 3 favoured large classes, and 3 showed no differences. The other 11 studies were evaluated by an analysis of desirable class activities and teacher practices, and all found small classes superior in these respects.

Several large-scale surveys in the U.K. have not resolved the issue. The Scottish Mental Survey of 70,000 eleven-year-olds (Maxwell, 1958) found no differences in test scores in relation to pupil/teacher ratio. A longitudinal study of a national sample of 5,000 U.K. children found that children in smaller classes performed slightly better, but nearly all these children were in private schools, so the results were not conclusive. Morris studied the reading progress of English children aged 7–11 years in 51 schools, and found that "schools with an unfavourable pupil/teacher ratio returned higher scores on the whole than those with smaller classes!" She added that the large classes usually had other favourable circumstances, however. Peaker in 1967 found the same pattern in his national survey for the Plowden Report, and made the same qualifying comments.

More recently the problem has been tackled internationally by the surveys of the International Association for the Evaluation of Educational Achievement (IEA). In the mathematics survey, students in countries with larger classes scored higher at the 13-year-old level, but the reverse was true at the 17–18 year level. It was suggested in explanation that, at the lower level, backward classes were reduced in size to assist their progress, and good teachers were given larger classes because they were able to cope with them. At the higher level, there would seem to be a tendency for the more advanced students to have smaller classes.

The IEA surveys in reading, science and literature, published in 1973, showed some relationships between class size and achievement but these were reduced to zero once the effects of home background and type of school/programme were taken into account. It is clear, then, that the problem of class size and achievement is far more complex than appears at first sight. Many other factors must be taken into account, and only sound experimental research will allow these to be fully tested.

The vast majority of research studies have used achievement of pupils on tests as the main criterion. An American study by Olson in 1971, however, looked at other "indicators of quality". Teachers in 18,500 primary and secondary classrooms were rated on four criteria of quality: individualization, interpersonal regard, group activity, and creativity. On these factors, smaller classes did produce significantly better results, and the relationship was consistent with classes ranging from 5 to 50 pupils.

However, as no measure of pupil achievement was used, no conclusion can be drawn about the importance of these 4 factors in promoting better learning amongst pupils.

. . . And Thickens: Factors Which Intervene in the Relationship Between Class Size and Pupil Performance

Several factors are believed to complicate the research studies and explain why children in small classes do not show up to better advantage. Some argue that when classes are homogeneous in ability they can be taught more efficiently in large groups, although there is no actual research which supports this view.

Again, small classes are more commonly found in rural areas and in older urban districts with bad social conditions. The children in such classes are less likely to perform as well as those from urban middle-class schools which frequently have large classes. Of course the opposite would be true of the small classes found in elite private schools. Slow learning students placed in small classes would also complicate the research picture.

Size of class also varies according to level of education. Smaller teacher/pupil ratios are found in the upper secondary school and universities, although the research has rarely supported this trend. Justification is found rather in the need for more space and more specialised facilities for the older, more advanced students.

Another complicating factor is the use of ancillary assistance in the classroom. English studies have shown that only about 43 per cent of the teacher's time is spent on lesson instruction. Therefore, if teacher aids were used, the teacher's efficiency would be increased, and the class size maintained at a higher level without loss of pupil achievement.

More important still is the behaviour and method of the teacher. Most would agree that different methods of teaching are best suited to different sized groups, and the more modern methods which focus on the way pupils learn rather than on the way teachers teach necessarily require smaller classes. Thus the lecture method can be used with equal efficiency for 200 students or for 10, but when used with large groups there can be little individual interaction between teacher and student. Even at primary and secondary levels, large groups of 100 or more are quite satisfactory where pupils are watching films or listening to illustrated talks. As soon as pupil participation is required, however, smaller classes become necessary, and with methods demanding close and continual interaction between individual pupils and their teacher, the smaller must the classes become to remain efficient. As one teacher put it, it is impossible to take an oral approach to French with classes of 40 pupils.

However, there is research in both England and America which shows that teachers do not automatically change from traditional lecture procedures when their class sizes are reduced. This fact would help explain why smaller classes do not show better in research studies. Improvement in pupil achievement will only follow from a reduction of class size if the teachers change their methods to take advantage of smaller groups, and if these changes are accompanied by genuine changes in the attitudes and beliefs of the teachers concerned.

Thus, the main reasons why smaller classes have not shown better achievement in research to date are:

(i) The continued use of traditional teaching methods in small classes.

(ii) The placing of low ability students in small classes, and vice versa.

(iii) The failure to take socio-economic and urban/rural background into account.

(iv) The concentration of research workers on a narrow range of curricular objectives.

(v) The use of class size rather than pupil/adult ratio, thus ignoring the flexible use of staff for varying size groups.

Our Heroes Arrive and Start Unravelling the Plot . . .

Gene Glass and Mary Lee Smith of the Laboratory of Educational Research, University of Colorado, wrote *Meta-Analysis of Research on the Relationship of Class Size and Achievement* in 1978. It reports on a new technique of integrating the results of previous research. Their conclusion: that average pupil achievement increases as class size decreases.

Their research began with a four month literature search which turned up 300 reports, articles, theses, etc. They found the Australian Education Department report of 1974 and the Ontario Ministry of Education report of 1975 particularly helpful. Only 77 of the documents could be used. These yielded 725 comparisons of achievement in different class sizes. The data came from studies of 900,000 pupils back to 1900. Included were studies of the effects of tutoring in very small groups of one, two, or a few pupils. Class size was defined as the number of pupils being taught by a single teacher, giving a Pupil:Teacher ratio (P/T) for a class of thirty taught by one teacher as 30, for a supplementary maths class of four with a teacher as 4.

The results of the 77 studies were all expressed in different scales. These had to be reduced to show what Glass and Smith call achievement advantages, and expressed in a common way. This score (called a "delta") was the mean achievement score for the smaller class in a study minus the mean for the larger class, the difference being divided by the within-group standard deviation. To illustrate, Class A has 10 pupils, Class B in the same study 20. The students have been given a test with 50 items. The mean for Class A is 35, the mean for B is 30. The standard deviation for Class A and Class B is 10. The delta for this hypothetical case is $(35-30) \div 10 = 0.5$.

Such calculations are easy when means and standard deviations, etc., are given. But many were not and Glass worked out ways of estimating them from other data given.

Of the 725 deltas calculated 60 per cent were positive, indicating that achievement was higher in smaller classes. The average was 0.9. Further analysis revealed two important interactions. The size of the difference depended on the size of the classes being compared. It also depended on the quality of the research design. Effects were stronger in studies having good design characteristics. Further mathematical analysis produced the now famous graph.

Imagine a typical pupil in a typical class of 40. He can be described as being better than 50 per cent of all the pupils of his class-level in a country-wide achievement test. The same pupil in a class of 20 would be better than 55 per cent of this group. In a class of 15 he would be better than 58 per cent. In a class of 10 he would be better than 65 percent. And being taught in a group of five he would improve his achievement to exceed that of 74 per cent of his class peers.

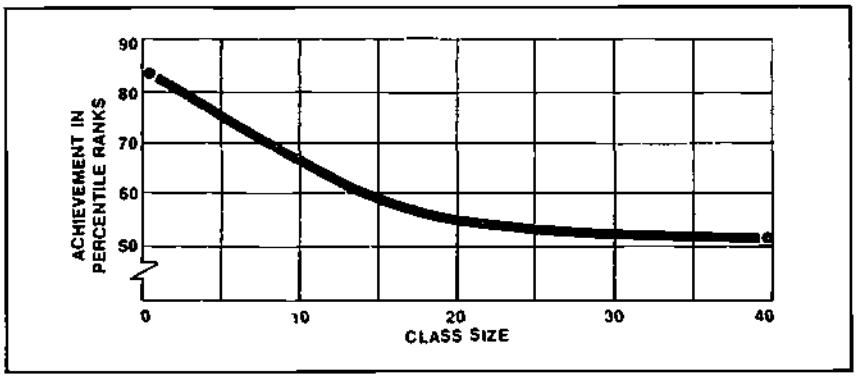
Another example: the average pupil in a class of 40 might be expected to gain one year's school knowledge in a year. In a class of 20 he could be expected to gain approximately 1 ½ year's knowledge. In a class of five he could be expected to gain 1.7 year's knowledge in a year.

In the figure, as in the examples, the curve starts to rise most dramatically when class size is reduced to below 20 pupils. Such improvement is not of course automatic and relies on the better use the teacher makes of the opportunities smaller classes give. Instances of substantially larger classes out-performing smaller classes have been recorded. Glass and Smith point out that "researchers must take account of what actually takes place in smaller classes: the instructional procedures used, the beliefs and capabilities of the teachers, the demonstrated abilities and backgrounds of the pupils, the subject matter, and the like. These ultimately determine whether the potential for increased learning that smaller classes create will be realised".

Glass and Smith found further that the relationship between class size and achievement was slightly stronger in secondary school than in primary school, but not strong enough to lead to the conclusion that primary pupils would profit less than secondary pupils if class sizes were reduced. There also appeared to be no difference in the results for reading, mathematics, and language.

The overall difference in results between the well controlled and the poorly controlled studies was dramatic. The curve for the poorly designed studies was almost flat, indicating, at best, a very small advantage to smaller classes. There was no correlation between class size and achievement in studies done before 1940.

How does one judge the importance of the differences shown in the figure? With such a small decrease in achievement between class sizes of 30 and 40 could it be argued that in these straightened times a return to 40 pupils per teacher would do little harm? Glass and Smith are continuing with their work and are now relating class size to classroom practice, and field studies are being conducted to see what changes take place in classrooms when they are reduced in size. In two schools four infant classes, each with a single teacher will be observed for half a year and then extra teachers will join the staff, the classes will be made smaller and the classes observed for the second half of the year. Such research as this will perhaps show why the increases are so dramatic as class size decreases below 20. However, in the meantime, it is



clear that increasing class size to save money will cause pupils to learn less. The effects such moves would have on teacher morale can be best spelled out by teachers' unions.

Would it be economic to increase achievement by reducing class size dramatically — to say, below 15? Already we reduce class sizes for some pupils, for example, backward readers, and reading recovery programmes report that the gains children make are worth the expense. The employment of less expensive paraprofessional or ancillary staff can free teachers for professional duties (that is, for teaching, which is often pushed aside by administrative chores). Grouping and individualised programmes are also methods of cutting down the instructional group size, if not the class. Staggered hours so that children "glide" in, some coming early for reading instruction and going early, others arriving later for school and reading, and going later, would also cut down the size of the instructional group.

In the end, however, weighing up outcomes against costs is a question of values. Glass shows a clear relationship between class size and achievement. What value does society, and therefore the school, put upon a particular magnitude of improved achievement, and upon those values the school can give that are not measurable?

After the Cheering Has Died Down: Our Heroes Methods Are Questioned

In essence Glass and Smith's techniques boil down to an attempt to quantify — to put into numbers — a review of all the available literature on the subject. The danger of this is that it lends the results an appearance of more accuracy and veracity than they may deserve. There are two important and related points which need to be made and which should influence any interpretation of the findings.

(a) There are very few studies which contrast small classes (i.e., 1–10) with classes a little bigger than themselves (i.e., 2–16) thus the graph as drawn may reflect the influence of a very small number of results.

(b) A substantial proportion of the effect produced by the analysis may result from the comparatively small number of studies which contrast individual instruction with larger classes. The inclusion of these studies may give the Glass Smith graph an unwarrantedly sharp bend and exaggerate the real gains to be made from reducing class size where, for us, it is most realistic, e.g., from 30 to 25. Nevertheless, the graph would be expected to follow the same shape, even if the one-to-one classes were omitted. However, the slope would be less steep at moderate class sizes and apparent gains from reducing class size much reduced. The slope at very small class sizes would be, however, much steeper.

Redrawing the graph to take these considerations into account could possibly reveal that Glass and Smith's results have tended to *over*estimate the gains in achievement resulting from a reduction in class size from 30 to 20, but have *under*estimated the gains resulting from individualised instruction.

To fill out (and perhaps to justify), the Glass and Smith predictions of large gains in very small classes many more pieces of research on the differences between classes with few members are needed, for example, classes with three, four, or five members contrasted with classes of six, seven, and up to 16 members.

Overall the Glass and Smith finding of increased performance from smaller classes holds. But the effect is probably less than they show. Alas, the fate of our heroes is, in the meantime, to have their importance, like their graph, lowered.

Credits:

The first part of this article was written by Warwick Elley, now Reader in Education, University of the South Pacific, Fiji, for New Zealand *set* in 1975, summarising the work of Douglas Pigeon in 1973 for the OECD, published by them as "Class Size as a Factor of Pupil Performance: A Policy Analysis".

The second part, *Our Heroes Arride . . .*, is a summary of two articles: the first by Gene Glass, Leonard Cahen, Mary Lee Smith, and Nikola Filby, "Class Size and Learning — new interpretation of the research literature" in *Today's Education*, April-May 1979; and "The Class Size/Achievement Issue: New Evidence and a Research Plan" in *Phi Delta Kappan*, March 1979, by Leonard Cahen and Nikola Filby.

The third part, *After the Cheering . . .*, is by Peter Jackson of the Test Development Unit of the New Zealand Council for Educational Research and is based on examination of Glass, G.V., and Smith, M.L., *Meta-Analysis of Research on the Relationship of Class Size and Achievement*, Boulder, Laboratory of Educational Research, University of Colorado, 1978.

A report drafted by Mudina Campbell of the New Zealand Department of Education Research and Statistics Division was of help to the editor.

References:

The quotation about large classes being exhausting, etc., is from Fleming, C.M., "Class Size as a Variable in the Teaching Situation" in *Educational Research*, 1, 2, 1959.

Blake's 1954 review of class size research is found in Blake, H.E., "Class Size: A Summary of Selected Studies in Elementary and Secondary Public Schools", Ph.D. dissertation, New York, Teachers' College, Columbia, 1954.

The Scottish Mental Survey can be found written up in Maxwell, J., *Educational and Other Aspects of the 1947 Scottish Mental Survey*, London, U.L.P., 1958.

The longitudinal study of 5 000 U.K. children can be found in Douglas, J.W.B., *The Home and the School*, London, McGibbon and Kee, 1964.

Morris, who found higher scores where the teacher/pupil ratio was unfavourable, can be read in Morris, J., *Reading in the Primary School*, London, Newnes, 1959.

Peaker, G.F., wrote "The Regression Analysis of the National Survey" which is Appendix 4 in the Plowden report, *Children and the Primary Schools*, London, H.M.S.O., 1967.

The American study which looked at indicators of quality other than academic success was described in Olsen, N.M., "Ways to Achieve Quality in School Classrooms: Some Definitive Answers" in *Phi Delta Kappan*, 53, n., pp. 63–65, 1971.

The Australian report on class size is LaFleur, C.D., Sumner, R.J., and Whitton, E., *Class Size Survey, October 1974*, Canberra, Australian Government Publishing Service for the Department of Education, 1975.

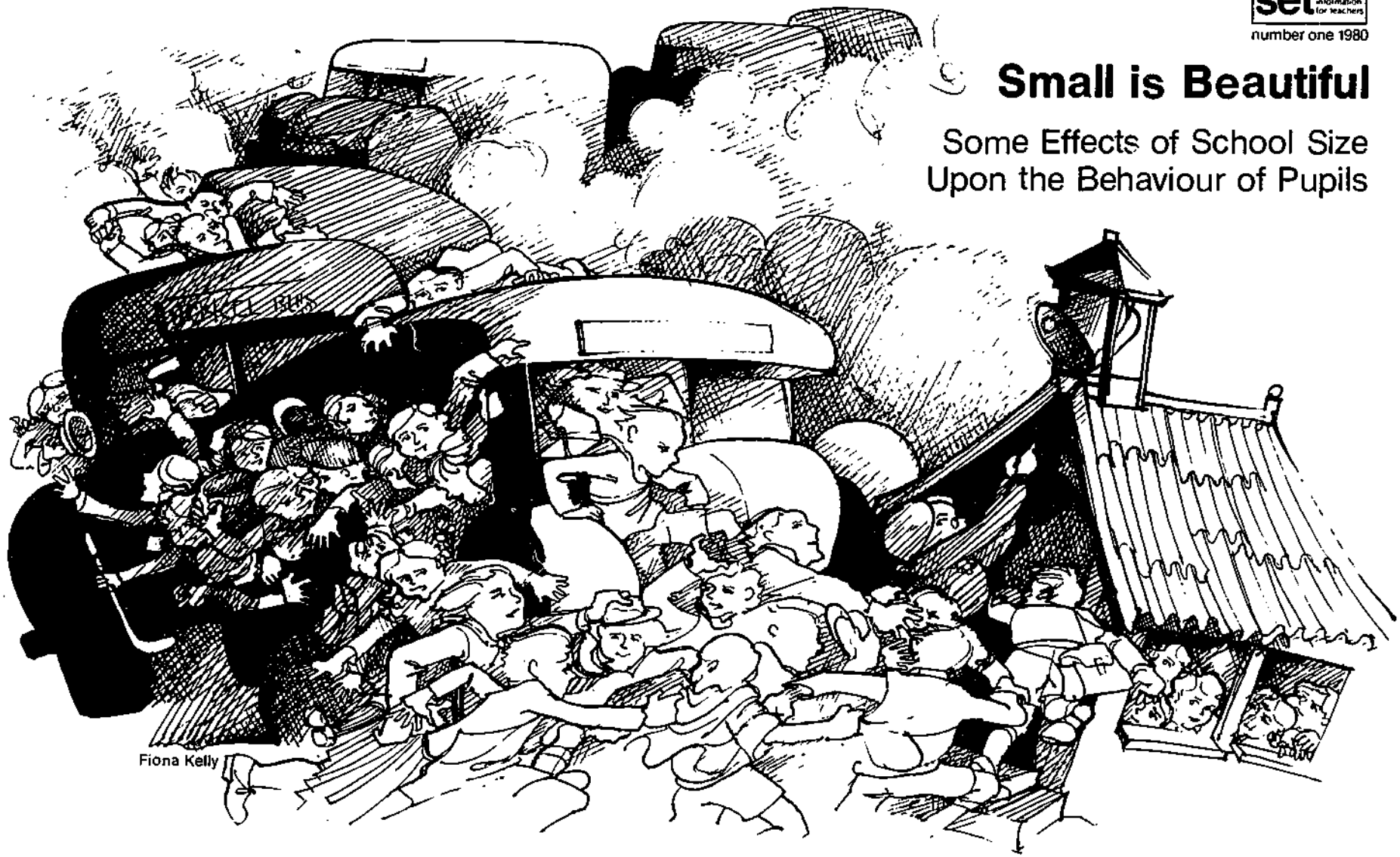
The Canadian report is Ryan, D.W., and Greenfield, T.B., *The Class Size Question: Development of Research Studies Related to the Effects of Class Size, Pupil:Adult and Pupil/Teacher Ratios*, Toronto, Ontario Ministry of Education, 1975.

Item 8

set research information for teachers
number one 1980

Small is Beautiful

Some Effects of School Size Upon the Behaviour of Pupils



Fiona Kelly

Small is Beaufiful: Some Effects of School Size on the Behaviour of Pupils

By Jack Campbell,
Professor of Education,
University of Queensland.

Introduction

Is the issue worth investigating? After all, numerous studies have shown that the home environment is a major influence on behaviour, and, within the school, more promising variables for study are teacher-pupil relationships, the curriculum, and specific teaching strategies. I concede at the outset that the size of a school which a child attends is unlikely to be a *major* influence on his behaviour, but it could be a very pervasive one — millions of children in Australia and New Zealand attend schools, and even a minor influence would, in total, be very substantial. Moreover, however minor its effects might be, school size, unlike certain aspects of home background, for example, could be controlled if need be.

The question of whether or not the issue *can* be investigated convincingly is more difficult to answer. Although many of our everyday actions are based on the assumption that our physical situation can affect us psychologically, it is notoriously difficult to present a convincing demonstration of this relationship. However, I want to try to proceed as far as possible in identifying effects which flow *almost inevitably* from school size and are not influenced to any great extent by other school factors or even the personalities of teachers or pupils. Sooner or later, of course, I will have to leave the shelter of allegedly inevitable relationships, and enter the stormy water of *probabilities*.

Characteristics of Large and Small Schools

The one absolutely sure thing is that, in comparison with large schools, small schools have fewer pupils! Both large and small schools can be further characterized by what they offer their pupils, for example, regular classroom lessons, special-occasion events (such as a visit to Parliament House), and extracurricular activities. In the jargon, these offerings are called 'behaviour settings'. Study after study has shown that although the bigger schools tend to have a bigger list of behaviour settings, the difference in number is often small, and the difference in variety is frequently nil. The smaller schools tend to offer the same kinds of things as do the bigger ones, but there might not be as many instances within each variety. Thus in a recent study which some of my colleagues and I have just completed, large schools with enrolments above 700 offered their Year 7 pupils an average of 60 behaviour settings (there were 6 varieties) whereas small schools with enrolments below 400 offered their Year 7 pupils an average of 50 behaviour settings within the same number of varieties (6). On the one hand, teachers in the small schools appear to make special efforts to ensure that their pupils are not disadvantaged in comparison with their large-school counterparts, and, on the other, large schools appear to be unable to increase offerings in line with increasing enrolments.

Some 'First-Order' Consequences

The first 'almost inevitable' consequence is that pupils in small schools participate in more activities, a greater number of key positions, and a wider range of activities than pupils in large schools. There is nothing mysterious, or even psychological, about this; it is little more than a matter of arithmetic. With fewer pupils available to 'man' the behaviour settings, it follows that greater involvement will be required from those who are there. In a very small school, every pupil might be called upon to play a key role in every activity, whereas in a very large school many pupils are likely to be redundant.

A second first-order consequence is that pupils in small schools will experience greater challenge and greater support than pupils in large schools. Each behaviour setting contains a system of controls designed to maintain it intact and functioning at a stable or near-stable level. Thus when pupils are in abundance, the less competent are likely to be relegated to onlooker roles, but, when pupils are in short supply, some are likely to be involved above their levels of competence. In order to ensure that the setting functions adequately, these latter will be given supportive guidance.

The picture that has been built up, then, is of pupils in small schools participating at higher levels in a greater number and variety of settings whose psychological climates are more challenging and supportive. Every study of school size, irrespective of teachers, pupils, or other variables, shows this to be so and the reason is that these first-order consequences are based on how behaviour settings function under conditions of 'over-manning', 'optimal-manning', and 'under-manning'. In a sense, they are *processes* rather than *outcomes*.

Some people may be impressed by these first-order differences in behaviour, but others are inclined to say 'O.K., but so what? Do these differences in behaviour within schools of different sizes persist and/or influence other more significant behaviour?'

Persistence of First-Order Behaviour

Two of our recent studies suggest that there is at least some persistence of behaviour generated by large and small schools. In the first study we looked at one small school which had been established, a year before the data were gathered, by drawing pupils from a very large school (1600 +). When we presented the computer with all the data in the study and asked it to cluster the schools according to the behaviour profiles of their

pupils, it made an almost perfect separation of small-school and large-school groups — but placed this newly-established small school with the large schools. Obviously the pupils from this particular small school were still, after one year, displaying the first-order behaviours which are typical of large-school pupils.

In the second study attention was focused on the behaviour of pupils who had attended small, rural schools for the first ten years of their schooling, but had then had to move to large secondary schools for Years 11 and 12. Typically, a small group of around ten students from a 'decapitated' school had joined about one hundred and fifty others who had been together for at least three years. In these circumstances, one might not expect the ten commuters to make a major splash in the social and extracurricular pools of their new school, but we are finding that they do. By very much more than chance would suggest, the Year 12 prefects and captains are being drawn from the groups of pupils who had spent their first ten years in small schools elsewhere. Before coming to their new large schools, they had, of course, been the most senior pupils in their small schools, and, almost inevitably, had occupied positions of leadership.

Influence of First-Order Behaviour on Other Behaviour

If one assumes that processes of development are basically similar for most people, and that experiences impinge similarly upon most, it is possible to make some predictions from the first-order behaviours to second-order ones which are 'probabilistic'. An important prediction is that:

> pupils who participate at higher levels in a greater number and variety of settings whose psychological climates are more challenging and supportive (in other words, small-school pupils) will develop: 1. a *stronger sense of cohesion*, 2. *a more favourable attitude towards school*, 3. *a greater concern for persons*.

Working with Year 7 pupils who had spent at least five years in their small or large primary schools, we have confirmed these predictions. After account has been taken of all other factors, it appears that between 10 and 20 per cent of the variance in scores on those three outcomes is explained by the first-order consequences of school size. Perhaps it was these second-order effects (stronger sense of cohesion, a more favourable attitude towards school, and a greater concern for persons) which led to the ex-small-school pupils quickly moving into positions of peer and school leadership within the large schools to which they moved after Year 10.

How Significant are the Findings?

Our findings are that:

1. in comparison with large-school pupils, small-school ones participate at higher levels in a greater number and variety of settings whose psychological climates are more challenging and supportive;
2. this orientation towards involvement persists, at least for a short time, when pupils enter contrasting settings;
3. the first-order effects appear to generate second-order ones pertaining to sense of cohesion, attitude towards school, and concern for persons.

I suspect that the significance ascribed to these findings will depend largely upon the educational values which one holds. Industrially-oriented readers will note that I have presented no evidence to hint that school size effect work-force skills. Nor have I presented any about different knowledge in the basic subjects of reading, spelling, writing, and arithmetic. These readers are likely to conclude from this that the economies of scale, which bigger schools are said to have, might as well be enjoyed. Similarly, academically-oriented readers are likely to dismiss the findings as insignificant because they say nothing about either the processes or products of thinking. I respect the value stances of these two groups of readers, and agree that the educational outcomes which they hold dear are important. Our young school leavers should be able to earn a living and contribute to the economic growth of our countries. These are, *of course*, important assets both for the individual and for society. Young people should, too (particularly as we begin to move into a post-industrial era), possess the intellectual skills needed to gather, store, and retrieve information. These are all necessary attributes. But they are not sufficient for living in this modern world. In addition, our young people need the motivation which will assist them to build a more humane world and overcome the challenges of war, poverty, pollution, ignorance, disease, and diversity of values within and between nations. I believe that this motivation is likely to be associated with a concern for people.

A favourable attitude towards school is no less important. As Bill Renwick has said, 'The extent to which each person is equipped to respond positively and confidently to the educational challenges that will face him as an adult will greatly depend on the attitudes engendered during his initial encounter with the education system.' In the interests of continuing motivation, which in our changing society is emerging as a more important educational outcome than specific items of information, we need to foster a favourable attitude towards school among all our pupils.

Few would wish to claim that by changing the schools we can *guarantee* the creation of a significantly more humane world, and the fostering of continuing motivation in the young people. Nevertheless, if, by changing the size of schools, some improvements can be effected in these areas without loss to other legitimate goals, change is warranted.

What should be done?

Three common-sense solutions suggest themselves:

(a) expand resources and facilities in large schools;
(b) create a number of semi-autonomous schools within each of the large ones; and
(c) ensure that the schools are kept small.

The first of these solutions is frequently difficult to implement — there is a limit to how many football fields or netball courts one can place on a school site — and is unlikely to work, anyway. Better resources and facilities lead to better educational experiences only when they actually impinge upon the pupils, and all of the research evidence on school size suggests that higher density generates lower involvement on the part of the pupils *despite the resources made available*.

The second solution sounds plausible, but in educational matters we have all been fooled before by plausible-sounding, 'common-sense' solutions. A recent study by Elizabeth Campbell suggests that 'vertical-unit' schools are much the same as others of the same *total* size in matters of 'flexible learning structures'. It appears that the negative relationship between school size and flexible learning structures is deeply based and difficult, if not impossible, to avoid. Nevertheless, there should continue to be active experimentation with different structures within the large schools which currently exist. The *whanau-house* concept in New Zealand seems particularly promising.

The third solution, which involves the creation, or maintenance, of small schools, appears, at the moment, to be most worthy of support. Research does not provide data on what the optimum size of schools is, but it suggests that when schools have enrolments above 700 it is difficult to ensure that the pupils participate at high levels in warm and challenging learning environments. At the other end of the scale, there is a hint that very small schools within communities with a low rating on a socio-economic scale frequently require compensatory assistance.

It might be claimed that I am flogging a dead horse, for most of what is contained in this article has been acknowledged by teachers for a long time. Despite this acknowledgement, however, one of the most striking trends in education during the last half-century, when populations have been rapidly increasing, has been a reduction in the number of schools. This reduction, of course, has been accompanied by an increase in the average enrolments. It is a depressing experience even to visit some of our large inner-suburb schools and see up to 2,000 pupils being packed into areas that contain only a ramshackle set of buildings and no playing areas. What must it be like to *learn* there? At no other stage in their lives will persons occupy such densely-populated, constrained settings for such extended periods of time. The industrial arguments of scale *still* dominate educational thinking in this matter: big schools provide better resources; better resources mean better experiences for pupils. Neither of these arguments is sustainable; it is time to announce that the Emperor has no clothes.

Notes

Characteristics of Large and Small Schools

The recent study mentioned is by Campbell, W.J., Cotterell, J.L., Robinson, N.M., and Sadler, D.R., *Effects of School Size on Some Aspects of Personality*. Report to the Education and Research and Development Committee, Canberra, 1979.

Persistance of First-order Behaviour

The first study mentioned is that just above. The second study is by Campbell, W.J., and Robinson, N.M., *Some Effects of Consolidation Upon the Quality, Efficiency and Economy of Education in Rural Communities*, Report being prepared for the Education Research and Development Committee, Canberra, 1980.

What Should be Done?

The study on vertical-unit schools is by Campbell, Elizabeth M., *The Relationship of School Learning Structures to Some Reactions of Students*, Report to the Education Research and Development Committee, Canberra, 1979.

How Significant are the Findings?

Bill Renwick is Director-General of the New Zealand Department of Education. His remarks on attitudes to school are from Renwick, W.L., *Social Change and the Objectives of Secondary Education* in Codd, J.A., and Hermansson, G.L. (Eds.), *Directions in New Zealand Secondary Education*, Auckland, Hodder and Stoughton, 1976.

set research information for teachers
number one 1980

Projecting Education Expenditure

By Ian McGill,
Department of Education,
and others.

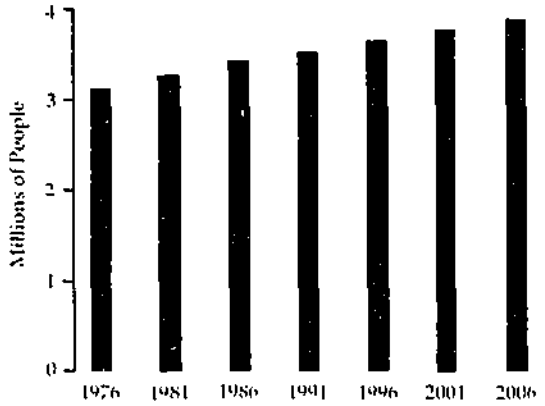
Fiona Kelly

In 1979/80 the government set aside $1,047,140,000 for education. That is 5.2% of our Gross Domestic Product. Spending on Education has been increasing but now fewer babies are being born, immigration has almost ceased, emigration is at an all time high, times are tougher economically, so how may education expect to fare?

Experts on population changes and finance can look into the future and show us our options reasonably clearly, but they cannot predict the future. Calculated guesses about school enrolments will, for example, help us invent possible ways of spending our taxes in the future, but they will not tell us how the money will in fact be spent. Everyone must decide which possible solution fulfils their hopes best and then tell the politicians; they have the job of finding the most acceptable solution.

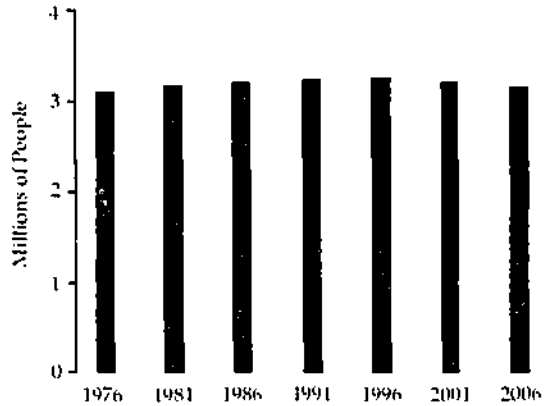
Demographers (population experts) like to keep their options open, so they present us with different models. For example, we can assume that the birthrate stays steady for some time to come and the effect, coupled with zero migration (immigrants and emigrants cancelling each other out) will give us the model of population growth seen in Chart 1 on the next page.

Chart 1 Total Population Model A — steady birthrate, zero migration



However, we could assume that the birthrate will continue to fall, and that we will lose 10 000 people per year. This will give us model D: Chart 2.

Chart 2 Population Model D — birthrate dropping, emigration continuing



There are models in between.

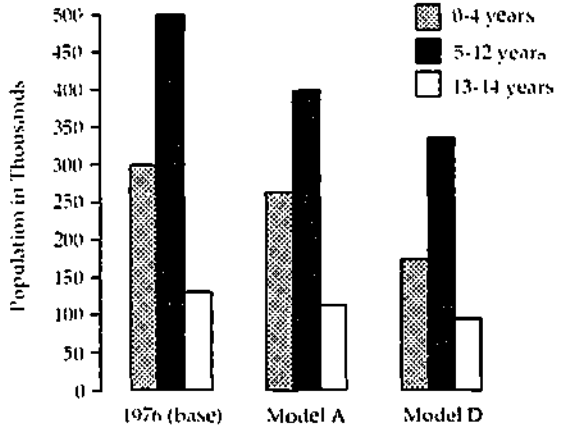
Since these projections were prepared actual emigration has been greater than assumed in D and in 1979 the population has fallen already. Is this decline a trend or a hiccup?

Even if it is assumed to be a hiccup from the point of view of educational planning, it is not so much the size of the total population that is of special interest but of certain age groups within that total:

1. pre-schoolers 0-4 years
2. primary schoolers 5-12
3. secondary schoolers 13-14 (compulsory)
4. secondary schoolers 15+ (voluntary)

If we compare the two models with our base year of 1976 we get chart 3.

Chart 3 Projected population in 2006



Model A shows a decrease in the size of the age groups of about 10% despite the steady rise in the population of the whole country.

Model D shows a decrease in the size of the age groups of about 40% although the total population is bigger than it was in 1976.

Most demographers suggest that Model D is the most plausible for planning purposes — they know that in the past they have consistently over-estimated and even D may be rather optimistic considering our 1978 decline. All other things being equal it implies a one-third reduction in our demand for maternity beds, paediatricians, teachers, schoolrooms, training college students, and so on.

All other things are not, of course, equal. Some sections of the community will be hit more than others; the effect on rural areas may be devastating. First, however, finance.

In 1978-9 $936,671,000 was spent on education under the title 'Vote Education' by the government and there were 1,006,864 students of one sort or another from Kindergarten to full-time University students and part-time adults at night school. The simple division sum gives us $930.29 per student but we should look deeper than that. After all each playcentre child costs the government about $82 per year and the Minister of Health recently gave the cost of training a doctor as $12,500 per year.

Education Expenditure

'Vote Education' is divided into seven descriptive programmes further divided into activities. A condensed version for three years is shown in chart 4.

Chart 4 Vote Education in millions of dollars

Programme		1971-2	1976-7	1979-80 (estimated)
I	Administration and General	7.6	13.3	20.7
II	Pre-school	2.4	8.9	15.6
III	Primary and Secondary	189.4	384.7	614.5
IV	Continuing	77.2	175.2	279.5
V	National Library	2.1	4.0	7.4
VI	Research	—	—	1.8
VII	Buildings	52.2	118.5	107.6
	TOTAL	331.0	704.7	1047.2

If we look more closely at primary and secondary school costs we get chart 5.

Chart 5 Primary and Secondary Education

Year	1971-2	1976-7	1979-80
Number of State Primary Pupils	468 267	475 113	467 500
Number of State Secondary Pupils	161 442	197 912	197 416
Total Pupils	629 709	673 025	664 916
State Primary Expenditure	$101.8m	$196.4m	$301.8m
State Secondary Expenditure	$ 66.1m	$131.4m	$215.3m
Joint Primary and Secondary Expenditure: on administration, special education, buildings, etc.	$ 51.6m	$112.8m	$145.6m
Total Expenditure	$219.5m	$440.6m	$662.7m
Cost per State Primary Pupil	$ 217	$ 413	$ 646
Cost per State Secondary Pupil	$ 409	$ 664	$ 1091
Average Joint Cost per State Pupil	$ 82	$ 168	$ 219
Average Cost per State Pupil	$ 349	$ 655	$ 997
State Primary, Secondary and Joint Costs: as a percentage of total cost of Vote Education	65.0%	62.5%	63.3%

Since 1971 direct expenditure per primary student has trebled and for secondary students increased about two and a half times. The increases have not been even 'across the board' — administration and special education have increased more than school buildings.

These figures must now be adjusted for inflation.

Chart 6 Cost per Student (in 1979-80 dollars)

Year	1971-2	1976-7	1979-80
Cost per State Primary Pupil	$ 573	$ 618	$ 646
Cost per State Secondary Pupil	$ 1079	$ 993	$ 1091
Average Joint Cost per State Pupil	$ 216	$ 251	$ 219
Average Cost per State Pupil*	$ 921	$ 980	$ 997

* Not simple averages from the figures above.

So in real terms, at most, there was an 8.3% increase between 1971 and 1980 (the highest average cost per student was in 1978-9: $1021. The small decrease of 2.4% from 1978-9 to 1979-80 was mainly due to a reduction in spending on buildings.).

The proportion of Vote Education going to Primary and Secondary Schools has not varied to any great degree.

The Future

Research in the USA and Canada has shown that when rolls fall, spending on education does not fall at the same pace. This has been taken into account in the following 'scenarios'. Primary and secondary school costs have been kept together for statistical reasons, despite the different pace at which their rolls will fall. The average cost per pupil from 1976 to 80 has been used for the calculations.

Chart 7 Four Scenarios for Educational Spending

		Change in Spending	Enrolments	Average Cost per Pupil
W	Severe cuts Falling enrolments	- 15%	low	$ 835
X	Moderate cuts Steady enrolments	- 5%	medium	$ 953
Y	No cuts Steady enrolments	0%	medium	$1003
Z	Spending up Enrolments up	+10%	high	$1103

These scenarios are further complicated if we take into account possible changes in school enrolments; for example, in times of economic hardship children may stay on at school trying to get higher qualifications, or enjoying the social life, rather than joining the unemployed. This gives us four scenarios worth considering.

Chart 8 The Four Scenarios' Effects on Future Spending on Primary and Secondary Education (in 1980 dollars)

	1980-1	1984-5	1990-1
W	$562.7m	$526.8m	$443.9m
X	$628.6m	$593.7m	$539.4m
Y	$661.2m	$624.9m	$567.7m
Z	$728.0m	$696.4m	$649.8m

The most surprising result is in scenario Z where an increase in spending per pupil of 10% and a high projection of roll numbers still results in a real reduction in spending on education of 2% by 1991.

But before we point out this automatic decrease in education spending to those who wish to cut the education vote there are a lot of other factors to be remembered.

1. Total school rolls decrease in all these scenarios but the mix does not remain constant. For example in scenario

W in 1980 secondary schools are 30% of the total enrolment, in 1988 they are 36% of the total school role. The cost of a secondary pupil is 60% more than a primary pupil so scenario W may be a further 3% more by 1988. If many of the extra secondary pupils are 6th and 7th formers more teachers will have to be employed, raising the costs further.

2. Any increase in the proportion of the educationally disadvantaged through social, economic, or other reasons may require further expenditure per student.

3. About 70% of educational expenditure is on people so salaries negotiated above the 'ordinary' inflationary moves — for example to attract scientists and mathematicians into secondary school teaching — will change the costs per pupil.

4. These figures make no allowance for the impact of roll decline on individual schools and teachers. Teachers' associations and parents will press for smaller classes, strong pressure groups will oppose the closing of small rural schools, early retirements will have to be paid for, busing pupils costs money and closed schools are not usually sellable, senior staff will probably be kept on rather than new inexperienced teachers employed and senior staff cost more, teachers' associations will negotiate salary protection measures as school rolls and gradings drop, and a significant part of school expenditure is not directly indexed to roll numbers.

5. Private schools integration costs money.

6. Unless the number of administrative and support staff drop proportionally with school rolls, costs per pupil will rise.

Implications

Staff

As some 60% of state school expenditure is on teachers, if falling roll numbers mean that salaries are renegotiated and put the salary bill up 2%, then in 1990 we may have the situation shown in Chart 9.

Chart 9 The Four Scenarios' Effects on Teaching in 1990

	Teachers' salaries	Average cost per teacher	Number of teachers	Number of pupils	Teacher/Pupil ratio	% Improvement in Pupil:Teacher ratio	% Reduction in the number of teachers since 1979
1979	$397.6m	$12,258	32 438	664 916	1:20.5		
1990							
W	$266.3m	$12,503	21 300	504 300	1:23.7	-16%	34%
X	$323.6m	$12,503	25 880	558 400	1:21.6	- 5%	20%
Y	$340.6m	$12,503	27 240	558 400	1:20.5	0	16%
Z	$389.9m	$12,503	31 180	583 700	1:18.7	+ 9%	4%

Competition for funds

Vote Education competes with other major central government activities such as 'Health', 'Transport and Communications', 'Social Services' and 'Development of Industry' for a share of government expenditure. Although Gross Domestic Product increased 22% in real terms over the last decade expenditure on education increased 32%. These rates of increase appear unlikely over the next ten years.

State school education must in turn compete with other educational activities: pre-school, university, technical and community education and, of considerable future importance, integrating private schools. In 1979-80 the primary and secondary private school appropriation is estimated at $27,650,000 for 80,480 private students or an average of $344 per student. By 1990 some 80% of private school students are expected to have integrated into the state system. This will create, on the basis of average costs, an additional demand in school education for some $28m under scenario W assumptions or $46m under scenario Z.

Resource Switching

An increase in Vote Education in real terms seems unlikely — it would involve higher taxes or a change in political priorities. An alternative, pleasing to those who get the increase, is to switch resources and expenditure from one area of education to another: the same sized cake but changes in the size of the slice you get. This has already happened, for example, as expenditure on teachers increased with improving staff/student ratios from 1978 to 80, and spending on school buildings declined.

Conclusion

There is no conclusion in this paper. Our governors need guidance in the face of falling rolls and economic hardship. A clear set of goals for educational expenditure is needed. What do we want education to do? William Taylor, Director of the University of London's Institute of Education, whose declaration that education must be accountable has delighted many, said at the same time, that the case for school education rests not just on competence but also on culture and citizenship.

Notes

The material in this article has been abstracted from three papers. The first two are by Peter Ramsay and James O'Neill of the University of Waikato and the third by Ian McGill of the Research and Statistics Division of the Department of Education.

All three papers include extensive bibliographies of sources and further reading.

O'Neill, C.J. and Ramsay, P.D.K., 'Educational Futures: The Implications of Demographic Change for Educational Policy' in *New Zealand Journal of Educational Studies*, Volume 14, No. 2, 1979.

O'Neill, C.J. and Ramsay, P.D.K., *Demographic Trends in New Zealand and their Relationship to Educational Policy Making*. Unpublished 1980.

McGill, Ian, *Projections of Education Expenditure*. Unpublished, 1980.

The two unpublished papers are available in photocopied form from NZCER.

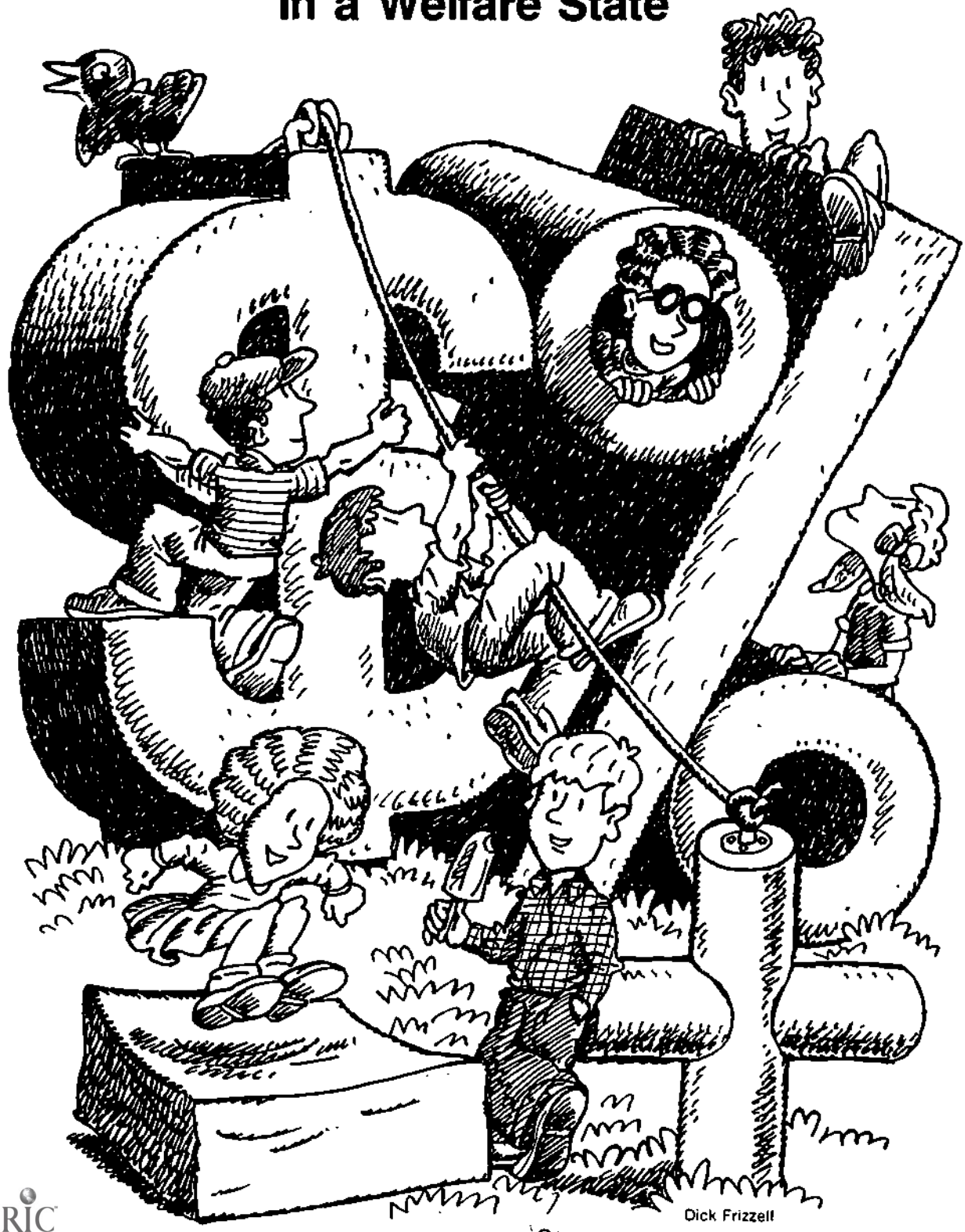
Ian McGill's paper (and the part of it used in this *set* item) expresses the views of the author and does not purport to represent the views of the Department of Education or any other official body.

Peter Ramsay and James O'Neill's papers contain carefully detailed descriptions of the effects of wholesale cuts in educational spending on rural schools and the quality of teaching, and draw conclusions about the opportunities for better quality education which falling rolls, and a good supply of teachers provide.

Item 4
(NZ)

The Economics of Children in a Welfare State

set research information for teachers
number one 1980



Dick Frizzell

The Economics of Children in a Welfare State

By Brian Easton, *University of Canterbury*

The Concept of Human Capital

In recent years economists have extended the notion of capital to cover similar phenomena in humans. For an economist a capital item generates a stream of future income. The traditional idea of capital included such items as machinery, buildings, installations, and the like. However, just as a capital good put to productive use generates income so does a human. With each person we can associate a future stream of income (but in some cases, such as parenting, it is generated in the informal economy and so does not appear as a financial transaction). As in the case of physical objects we can discount this income stream into a capital value. However, unlike objects, humans may not be bought and sold in the market (with some exceptions as in a contract to supply services), so the market value of human capital is notional. It can also be very high. The capital value of a person in his or her early twenties is likely to be twenty times his or her annual earnings.

One of the earliest uses of the human capital approach was to integrate education into economic analysis. Education was seen as a process by which the individual acquired more human capital, that is increased his future earnings. Since education involved resources and the student foregoing acquisition of human capital involves an investment process — that is the foregoing of consumption (or income) now, in order to generate higher consumption (or income) in the future. Economists were able to measure the return on this investment, which they found to be very large by normal standards. For instance in New Zealand later secondary school and university education gives a real return of 15 to 25 percent per annum (above the rate of inflation) to the individual and 5 to 10 percent per annum to the society.

Another use was in the aid programmes for developing countries where it was shown that good nutrition and health services for children and adults improved their potential to earn income. This suggested that it may be just as sensible to invest in human capital for foreign aid as it is to invest in physical capital. Thus the human capital model gives a justification for 'soft' aid programmes, that is, aid programmes directly assisting humans, as well as 'hard' aid programmes of more 'objective' items.

Investing in Children

We can extend the notion further by arguing that childrearing is also an investment programme. That is each society maintains and promotes its income in the future by investing in its children today. The investment becomes realised as the children grow up to adulthood and produce. Part of that production is retained by the adults that were the children (as is was by their parents) but part is commanded by society for the use of others, including social welfare provision for the elderly, sick, and disabled, and for further investment in the next generation of children.

This process of social investment in children which is realised in adulthood, is vital for the working of a modern welfare state. The ability to provide adequately for the retirement of its elderly is dependent upon the productivity capacity of its working adults. If some years before, during their working life, the elderly had invested wisely in the children who are today's adults, then, they provided for their retirement. Moreover it is not merely a matter of economic inputs and outputs. The investment activity places a moral obligation on today's adults to provide for the elderly who supported them in childhood, and a moral example for today's children who are tomorrow's adults, to provide for today's adults in their retirement.

Another important social consequence of the investment in children approach is its transformation of the role of parents. Because parenting is mainly in the informal part of the economy, (exceptions include child minding for money and teaching) the parent activity as an economic activity is easily neglected. However, from the point of view of investment in children, parental care is an integral economic activity to be measured by the income parents forego in order to care for children. By far the larger part of this foregone income is by mothers. The implication is that parenting in general, and mothering in particular, can be treated as an occupational activity. Perhaps it lies, in professional standing, between medicine and teaching at one end, and economics and politics at the other. The implications for women's life cycle and employment patterns are manifold, but would take us outside this discussion.

It has been possible to make some estimate of the size of the child investment programme in New Zealand. The total can be divided into broad magnitudes as follows. The maintenance costs of children in food, clothing, housing, recreation, etc., comes to around 12 percent of Gross National Product (G.N.P.) i.e. total economic output. Public service expenditure, mainly on education and health, comes to around a further 5 percent of G.N.P. And the costs of child care, calculated as parent's market earning potential foregone, is conservatively measured as 8 percent of G.N.P.

In total, investment in children comes to around a quarter of Gross National Product, or about the same magnitude as all the other investments combined. However, ordinary investment activity is more heterogeneous than child investment, so the childrearing industry is our largest single industry. Moreover it is a very productive industry. It appears that the rate of return on investment in children is about three to five times the rate of return in physical investment, although we get such a low return on our physical investment the ratio is not as impressive as it might appear.

What we do not know, is which facets of child rearing are most productive. It would be great if we could report that the return on breast feeding was *X* percent. on child beating was minus *Y* percent, and on fluoridation of the water supply was Z percent. Unfortunately, we do not know returns on specific activities. Until we do we must be cautious at advocating increased expenditure on children in the vague hope that this will increase the productivity of the industry.

Traditional capital theory was often illustrated by examples from the wine industry, and students were faced with considering the economic return on paying for the annual priestly blessing on the stored maturing wine. Similarly much of our expenditure on children may have as much effect on the maturing children.

One interesting result from my research is that New Zealand has been investing relatively heavily in human capital throughout the post war era. Like the wine merchant building up cellar stocks, much of this investment has not yet begun to generate cash income. The implication was that as the children entered adulthood in the 1980s New Zealand would have experienced a marked increase in its growth potential. However, it now seems likely that much of that human investment will be dissipated because of unemployment and outmigration.

Human Investment and Household Income

Having identified the size of the human investment the next step is to examine how it is financed, that is, who foregoes consumption today in order to create the human capital which produces the consumption of tomorrow. While the state, that is, the community collectively, contributes about a third of the resources required, the remainder, amounting to 16 percent of GNP, is provided by the parents.

Thus the cost of the human investment programme falls unevenly across the community, with a childless couple both of whom work, contributing relatively little even though they are on high incomes while the family with young children and only one adult in employment contributing a lot despite being on a lower income.

Three points can be made in terms of the earlier analysis. First the childless couple will expect its social security benefits during retirement and disability to be paid by others' children and for those children to honour the moral commitment not to expropriate (without compensation) the financial investment the couple have accumulated for their retirement. This implies that the childless, (and those with few children) have a reciprocal moral duty to support those with children (and those with many children).

Second, while the family with young children may be a one-income family, it is not a one-worker family since, although not in formal employment, the mother caring for young children is a worker contributing to social output.

Third, if the family was producing physical goods then it could finance the project by borrowing against the security of the income the goods would produce in the future. However, since we do not permit slavery, a financial institution cannot loan against the security of the future income of children. Thus a family cannot finance its human investment programme analogous to the financing of an ordinary firm.

This is a rather academic way of deriving the obvious point that the material standard of living of families with children is likely to be below the national average. Data based on the 1973/4 Household Expenditure Survey suggested that 63 percent of children and 57 percent of parents were in households with below median material standards of living.

That families are in this situation is not in dispute. Nor could there be dispute, given the preceding analysis. There has been a little dispute about the numbers of families below the poverty line. Leaving aside the erratic intervention of politicians, at least one of whom has argued that the problem is a 'poverty of intellect', thus perhaps demonstrating what he was talking about, the dispute has been over the definition and measurement of poverty.

Family Poverty in New Zealand

It is now broadly accepted that there are two sorts of poverty. Absolute poverty is when the individual is unable to sustain life and health. It is a phenomenon we associate with poor countries which may have widespread disease and inadequate nutrition.

However, within more wealthy countries, it has been recognised that while there may be little absolute poverty, the share some groups get of community consumption may be so low that they do not, in the words of the 1972 Royal Commission on Social Security, 'feel a sense of participation in, and belonging to, the community'. This is clearly a crucial problem, given the moral underpinning of the economics of children that we have referred to earlier. Such people are in relative poverty, and the Royal Commission on Social Security argued 'that there is a need to ensure within limitations which may be imposed by physical or other disabilities that everyone is able to enjoy a standard of living much like that of the rest of the community'.

Thus there are two poverty principles and by implication two poverty lines. Following traditional analysis we might call the absolute poverty line 'the breadline', in which case the relative poverty line might be called 'the meat line' since it is frequently associated with an inability to afford meat regularly at meals.

In practice the meatline is set as the material standard of living of a couple on the basic invalid social security benefit. The argument is that it is unlikely that the

benefit level is meant to be below the poverty line, and it is also unlikely that we mean it to be substantially above either. Moreover, there has been a bit of survey research to suggest that households below this income level do have difficulty maintaining an acceptable life style much like the rest of the community, such as regular meat meals.

If we adopt the social security level as the poverty line, then the evidence points to around a quarter of children and a fifth of parents are in households below this standard of living. This is by far the largest group of poor in the community. Moreover the bulk of poor families are not solo parent families and others living on social security benefits. The vast majority are primarily dependent upon the wage their father can earn in the market.

The research has gone a stage further in an attempt to elaborate a description of what determines household income, and therefore, household poverty. Quite a lot is known about various aspects of the mechanics including father's incomes, mother's workforce participation, the family life cycle, the impact of taxation and benefit, and some of the factors about family size. However, there have not yet been funds available to build up a model which integrates these facets together into a comprehensive description of household economic behaviour.

The Consequences of Family Poverty

We know less about the consequences of this poverty. In principle there are three major issues. First its existence undermines the implicit morality of the welfare state as we discussed earlier. If one group serving a worthwhile function in society is poorly off, then other groups which benefit from social support are likely to be treated as illiberally.

Second is the question of family morale. No matter how worthwhile a social function families may be carrying out, if they are not socially rewarded their psychological state may be weakened. The rewards need to be two-fold: income which keeps their standard of living much like the rest of the community, and genuine public respect for the role of childrearing. It seems likely that low morale can have real effects on childrearing performance. The tired, slightly hungry, slightly sick or tense parent is unlikely to manage the children well during the difficult developmental periods. The same phenomena can lead to marital tensions and even marriage breakdown when other difficulties already exist.

On the whole, family morale seems to be much better than we might expect from the treatment we give families. This is partly because its public image is suppressed, particularly in order to protect children. As one headmaster remarked:

> If there is a struggle in the home situation you might not see it reflected in the children... there are a lot of parents who have their priorities in the right place.

The children are the last to suffer.

On the other hand the housewives' boycott outburst of 1977 almost certainly reflected a pathetic response of demoralised mothers to the economic pressures on them.

The third consequence is the direct effect of poverty on childrearing. The previous two paragraphs indicate that deterioration in parental morale can result in less effective child rearing. It is also evident that poverty can result in the child's nutritional standards, comfort, or health being sufficiently below the acceptable level to result in poor performance, at school for instance. Moreover, an overcrowded, dangerous, or nonstimulating environment for a child may result in poor development and even injury. A final direct effect would be where poverty leads to lower opportunity for the children. There is evidence this applies to both health care and education.

On the whole there is very little systematic evidence to support suspicions that family poverty is causing markedly inferiour child development in New Zealand. But this is a conclusion of agnostism rather than atheism. It is an argument for minimising family poverty until we have evidence that poverty does not have these effects, rather than leaving family incomes alone until the research evidence is collected (knowing full well that there would be no urgency to do the research). Moreover whatever the conclusion of the research, the effect of the existence of family poverty undermining the welfare state cannot be ignored.

Recent Trends in Family Income Maintenance

The evidence for family poverty was becoming apparent by the mid 1970s. It involved the salutory lesson that the past solution for maintaining family incomes had failed to eliminate poverty. This approach, which appears from the beginning of the welfare state in the 1890s, was to expect the wage system to provide adequate support for intact families without disability. (There was special provision for other sorts of families, commencing with Workers Compensation in 1900.) In 1914 a small income tax concession for children was introduced. In 1927 there was the beginnings of the family benefit. The period of the first Labour Government saw the beginnings of systematic assistance with health and housing. ('Free' education began by 1877.)

The approach failed partly because it lacked coherence, but perhaps more fundamentally because any family policy which is based on the labour-earnings structure is bound to fail since the worker's earning capacity has little relation to his family's needs.

Since family poverty became evident the approach has remained ad hoc, but more attention has been paid to modifying the income tax structure in favour of children.

In 1976 the young family tax rebate was introduced for families with a child below the age of 5. Currently it is $9.00 per week but abateable after the first $150 per week of income of the principle earner (father).

In 1977 the single income family tax rebate for families with a child below the age of 12 was introduced. Currently it is $5.00 per week, but abateable after the first $25 per week of the income of the secondary earner (mother).

In 1978 the tax scales were changed to levy more tax on the secondary earner, and hence lower tax rates on the principle earner. This change benefited one income families at the expense of one and a bit income families.

These changes were reinforced from October 1979, and the family benefit was doubled. It seems likely that from October 1979 the state will be contributing more to the financial support, measured in real terms, of intact families without disability, than at any time in the last 20 years and perhaps longer.

Nevertheless, such ad hocery is not likely to succeed unless it is very expensive. Some comprehensive approach to income maintenance for families is necessary, including the institution of tax credits (or negative income tax) and the integration of subsidies for child care into the system. It should also be possible to include non standard family situations such as solo parents and disability into the system.

The task of constructing a comprehensive income maintenance scheme is not an easy one. Besides empirical evidence it requires a sound analysis, of which the economics of children is likely to be part, and a social philosophy.

The Family as a Production Unit

Another policy area which has developed out of the economics of children arises from the closer attention given to the family as a production unit. If child rearing is an industry, perhaps our largest and most important industry, then the families are its firms. It is a natural step to considering how we may increase each firm's productivity.

The effect then is to shift attention from the child rearing activities which occur in the formal part of the economy, such as in hospitals, schools and children's homes, and are supervised by paid workers such as medics, teachers, and social workers, to the childrearing which occurs in the informal part of the economy of households supervised by unpaid parents.

Once the point is made it is evident that by far the most childrearing (and the most important childrearing) occurs in this informal area. Yet the most resources and the most attention is given to the formal areas.For instance in a total public education budget of one billion dollars, the actual amounts spent on the education parents give to their children, including parent education, is less than 0.1 percent of the total. The implication is that we can substantially improve childrearing by greater concentration on the family, not because it is more efficient *per se*, but because we are doing so little that we must be able to do better at little cost.

Economics has little to say on what measures should be taken, except to reiterate the importance of an adequate economic environment in which the family should operate. It is for other disciplines to identify other measures. Despite our lack of research of the intricacies involved in family behaviour, there is sufficient evidence of major differences between optimal behaviour and practice in a number of areas to suggest effective measures based upon improving the public's understanding of parental responsibilities and activities.

For instance, the children of young mothers are at greater risk. Yet there is no public education programme among adolescents to discourage teenage births. We have become so obsessed with the physical details of contraception, that we neglect the psychological and social aspects of family planning. Spacing between children is another area where public understanding could be usefully brought more into line with scientific knowledge.

There are also vast gaps between parental understanding of children's health, nutrition and behaviour, and the range of practices considered optimal by expert opinion. This is not the result of deliberate ignorance on the part of most parents who are probably keener to learn than the experts are to teach. The problem is the creation of adequately funded institutional means to enable parents to obtain the education they seek. To do this may well involve a revolution in our approach in the formal economy. Perhaps schools should be offering courses to parents, on fully paid bursaries, and leaving more of the teaching of children to the parents.

One way of summarising this approach is the recognition of the 'pre-primary care' area for child development. Our tendency has been to think of the first institution, be it school, medical centre or social worker, as the point where care begins. Put this way the view is absurd, but an astonishing level of advocation and policy implicitly assumes this acknowledged absurd view. For instance it is proper to be concerned with those activities labelled as 'preventative', but, that label reflects a view from the institution/professional world, rather than the family.

One of the unfortunate aspects of the welfare state has been to shift power, responsibility, and resources away from the pre-primary care area, that is the family, to the institutions and professions. A reversal of the shift is required both for human liberty and for economic efficiency.

It is very easy to prepare authoritarian measures to meet some problem area in child development, but the authoritarians might do well to reflect on the recently reported treatment of children in state social welfare

homes. Obviously some family units are far from ideal. Apparently the alternative may be little better or even worse. And the fact there are a few failures does not justify unnecessary impositions on all failures (whether we are talking of families, doctors, teachers, or social workers).

Directions

One interesting feature of the economics of children is its tendency to analyse children in typical, although diverse circumstances, rather than to concentrate on atypical circumstances such as solo parent families, ex nuptial births, child abuse, or parental and child delinquincy. Research on pathological situations is necessary, but it can result in an over preoccupation with such situation at the expense of the typical situation. Too often new social policy in New Zealand is proposed by psychiatrists, social workers and politicians responding to the pathological cases they meet. The result can be the tail wagging the dog.

It may be that the economic approach is going to the other extreme, subverting the richness and complexity of the psychological, social and economic phenomena we call family life into a simple economic model. Obviously there is such a danger.

Much policy making, particularly in the social and industrial areas, effects behaviour within the informal economy. To neglect such behaviour can easily result in the policy prescriptions being inefficient, ineffective, and even having the opposite effect to that intended.

Economists need a theory of the informal economy. If such a theory is also beneficial for other groups involved in the informal economy or in developing policy therein, then we have an example of the advantages of co-operative enterprise.

Notes

The following is a list of publications which generally enlarge upon the points made in the text. Some of the articles listed are available in mimeograph from the author, but it may be necessary to make a charge. Other people's research used in the text is identified in these primary sources.

The Needs Index (mimeo, 1973)
Family Incomes (mimeo, 1973)
Poverty in New Zealand (mimeo, 1973)
"Staying on At School" *NZ Monthly Review* (Nov. 1973)
"The Welfare State: Principles and Future", *New Zealand Politics:A Reader* ed S. Levine (Cheshire, 1975)
"The New Zealand Housing Market" *New Zealand Economic Papers*, 1976
"Poverty in New Zealand: Estimates and Reflections" *Political Science* (Aug. 1976)
"The Economic Lifecycle of the Modern New Zealand Family" *Australian and N.Z. Journal of Sociology* (Feb. 1977)
"Income Tax on Employee Earnings" *Quarterly Predictions* (March 1978)
"The Economics of Children" *Impact: The Burden of Choice* (Akld Principles Association, 1978)
"The Married Women's Property Act and the Domestic Purposes Benefit" *Woman Studies Association Newsletter* (June 1978)
"Income and Fertility in New Zealand: A Cross-Sectional Study" *The Proceedings of the Annual Conference of the NZ Demographic Society 1978*
"The Rise of Family Policy" (mimeo, 1978)
"The Economic Needs of Young Children" *Parents Centre Bulletin* (Spring 1978)
"Population and the Economy" *New Zealand Population* and C.J. O'Neill and W. Neville (Longmans Paul 1979)
"The Family Costs of Children" *The Proceedings of the Annual Conference of the NZ Demographic Society 1979*
"The Economics and Politics of New Zealand Poverty" *Poverty in Godzone?* ed G.W. Bryant (1979)
"New Directions in Health and Economic Management" *N.Z. Medical Journal* (forthcoming)
"Household Equivalence Scales" *New Zealand Statistician* (forthcoming)
Social Security in the Seventies: Policy Formation and Failure (University of Canterbury, forthcoming)
Social Policy Analysis in a Welfare State (Allen and Unwin, forthcoming)
"Some Household Expenditure Functions" (with C.M. Low and S.F. Connor, forthcoming).

Brian Easton is a lecturer in Economics at the University of Canterbury. This article is taken from a talk given to the Early Childhood Care and Development Convention, Christchurch, 1979.

set research information for teachers
number one 1980

Teacher Expectations and Classroom Behaviour

By Alison St. George
Massey University

Impressions of Other People

People are always making judgements about people. It is hard to imagine how we would get along with other people without our perceptions of them as a guide. The ideas we have about others influence the way we behave toward them and the way they respond to us. There are several important points here:

a. We are active perceivers, selecting, organizing, and structuring what we notice about people.
b. The situation and who we are as well as characteristics of the person affect what we see.
c. We make inferences and attribute characteristics to the person on the basis of immediate cues. The attributes may concern temporary states (e.g. he is bored) or more permanent characteristics (e.g. she is intelligent). Many of these characteristics cannot be observed directly.
d. We also respond in an emotional way, liking, being interested in, having sympathy for, rejecting, and so on.
e. Expectations are generated when we perceive other people and these are particularly important in face-to-face situations because we need to prepare for responding.

Teachers and Expectations

Any school-learning situation is essentially one of social interaction: teachers and pupils form impressions of each other, react emotionally, and anticipate attitudes and behaviour. Judgements and expectations guide behaviour, though those involved may not always be aware of it. Further, we do not always carefully select and weigh up evidence. Often we make haphazard judgements.

For teachers, forming expectations about their pupils is an integral part of their task. It is necessary in order to be able to plan work for pupils, to assess their progress, and to relate to them. Teachers are also required to make deliberate judgements and write them down for school records, for parents' information, or testimonials for employers' guidance. Thus the whole business of perceiving and judging other people has particular relevance for teachers and often has long lasting consequences for their pupils.

Teachers have expectations about their pupils' intellectual potential and academic achievement. However, they also face the task of social control, and their expectations about general classroom behaviour are also important. Expectations should not be confused with hopes, desires, and aspirations. What a teacher hopes a new class is like, and what he or she expects it will be like, are often not the same thing!

Expectations as Self-Fulfilling Prophecies

In the late 1960s it was suggested that a teacher's expectations for a pupil could be self-fulfilling. *Pygmalion in the Classroom*, published in 1968, caused considerable controversy. In it Rosenthal and Jacobson described a study in which they attempted to manipulate teachers' expectations to see if those expectations would be fulfilled. The findings were

interpreted as demonstrating that favourable expectations could lead to gains in pupil achievement and general ability. By implication, the study also seemed to support claims made by critics of the schools that the lower academic achievement of poor and minority group pupils was largely the result of the negative stereotypes and low expectations of their teachers.

Popular reports of the research made enthusiastic claims which were often exaggerated and the notion of teacher expectations acquired a mysterious, almost magical, quality as if expectations always became self-fulfilling. A teacher only had to make a prediction and it would come true. Not surprisingly, many teachers did not accept this. The study was also strongly criticized in academic circles for problems with the experimental design and methodology, the data, the way it was reported, and the overgeneralization of the results.

Since that controversial piece of research a large number of investigators have conducted studies of teacher expectations using a variety of methods. In some studies attempts have been made to manipulate expectations, for example, by providing false test information or fictional reports. In others, teachers own naturally formed expectations have been used. Some investigators have looked at teacher expectations and measures of pupil performance or self concept, while others have been more concerned with the way in which expectations may be communicated through interactions between teachers and pupils. The evidence now available supports the idea that teacher expectations can sometimes function as self-fulfilling prophecies but that this is by no means automatic or inevitable. For an expectation to be self-fulfilling it is not enough that it just exists. The person concerned must recognise what the expectation is and the teacher's behaviour must also be effective in moving the person in the expected direction.

A New Zealand Study

A recent study in New Zealand has investigated the perceptions and expectations of five Standard Three teachers and their pupils. No attempt was made to manipulate the teachers' expectations about their pupils. Rather, the teachers own expectations were studied. Early in the year they were required to rank their pupils' general ability. Using these rankings a high, middle, and low expectation group was selected from each class, 90 children in all (67 Pakeha, 20 Maori, 3 Pacific Islands), with equal numbers in each class and expectation group and equal numbers of boys and girls.

Perceptions and Expectations

The teachers were later asked to give further much more detailed perceptions of each child. Children for whom the teacher had high general expectations were seen as well off in:

reading
reaction to new work
independence (need for teacher guidance)
task concentration
use of English in conversation
home environment
interest in the world around
perseverance
participation
confidence
parent attitudes to school and education.

Not surprisingly pupils for whom the teachers held higher expectations were closer to the notion of an 'ideal' pupil (as identified in the research literature) than those in the middle and low expectation groups. They were perceived as having academic work skills and an eagerness to learn which makes it easier for the teacher to feel some success at teaching. They also tended to be seen as coming from a 'good' home background.

Teacher perceptions also varied with pupil ethnic group. Generally Pakehas were perceived as being closer to the 'ideal' pupil concept than the Maori and Pacific Island children.

Specifically, the teachers associated poor parent attitudes to school and education, less interest in the world around, a poorer use of conversational English, and a less stimulating home environment with Maori and Pacific Island pupils. Poorer reading skills, lower participation in class activities, a more negative reaction to new work, and a need for greater teacher guidance were also part of teachers' views of these children. To a lesser extent, the Polynesian pupils were also perceived as having a lower level of perseverance at and concentration on school work. The teachers expected less of the Maori and Pacific Island children, ranking them lower in general ability. Most were in the low expectation group.

The findings for teacher perceptions and expectations then do not look particularly encouraging as far as minority group children are concerned. It has been suggested by some that stereotyped expectations for Maori pupils may become self-fulfilling prophecies. Are teachers in New Zealand clearly prejudiced and as a consequence do they engage in discriminatory behaviour toward Maori and Pacific Island children in the classroom?

Expectations and Classroom Behaviour

How did the teachers behave toward their pupils? In order to study this we carried out classroom observations. Teacher and pupil behaviour was recorded whenever the teacher was interacting with a pupil but not when the interchange was with a group as a whole. Some of the situations were *public*, where the whole class was looking on. Others were *private*, for example, they were about a piece of work, behaviour, or classroom procedure concerning a particular pupil and not meant for the whole class.

Behaviour and ethnic group

There was no statistically significant difference between the way teachers interacted with Polynesian and Pakeha children. In fact there seemed to be an active effort on the part of teachers to give all pupils similar amounts of attention. While there may be a few teachers who actively discriminate against minority group children there is little evidence that this is generally the case. Certainly it was not seen in the study being described here.

Some interesting differences in classroom behaviour did emerge though when we looked at it in relation to what teachers expected of their pupils.

Public interactions

When children could respond in public there was no significant difference in the number of opportunities given to pupils in the high, middle, and low expectation groups. However, there was a difference in the type and complexity of questions asked. Sometimes teachers posed a question to a group and waited for hands to be raised before selecting a child to answer. In other instances teachers addressed a question directly to a particular child. When these two types of questions were compared it was found that high expectation pupils received most of the first 'open' type, while the low expectation pupils were asked most of the 'direct' questions.

The finding that teachers asked the low expectation group to answer fewer questions when they were posed to the class (and to which they expressed some willingness to respond) may be the result of the teachers' experience of the pupils' differences in knowledge and ability. But this trend is balanced by teachers addressing more questions directly to low expectation group pupils and selecting them to respond even when they have expressed no willingness to do so. The questions posed in this way were more likely to be relatively simple, requiring a short answer, rather than more complex process questions involving going through steps to solve a problem or explain a phenomenon. It may be that teachers have different reasons for asking public academic questions of children from the different expectation groups and that these reasons influence the complexity of the question asked. For example, in some cases the aim might be to see if the pupil understands the material, in others it might be to encourage participation.

What are teachers' reactions to pupils' public academic responses? Our findings were limited by our data. It was clear, however, that all teachers criticised pupils to a greater degree when they did not attempt a response than when they gave an answer which was wrong.

It may be that teachers view academic failure as more reprehensible when it is accompanied by an apparent lack of effort. However, frequent criticism for not responding does not seem likely to encourage pupils to respond in the future unless strategies for reaching solutions are offered and encouragement given. The informational value of criticism may be offset by the lower perceived warmth of the teacher. Further, the teachers as a group criticized the pupils for whom they held high expectations a lot less than either middle or low expectation group pupils — both in relation to academic work and to general classroom behaviour.

Private interactions

As with public responses, there was no significant difference between expectation groups in the overall number of private interactions they had with their teachers. However, a clear pattern did emerge when we asked, 'Who sought the interaction?' Children in the high expectation group created more private contacts with the teacher for themselves, whereas the teacher initiated the greatest number of private interactions with low expectation group pupils. Once again, there was an attempt by the teacher to compensate for the low rate of initiating interactions in the low expectation group and thus to reduce any inequalities in the amount of teacher-pupil interaction. This can be further understood by referring back to teacher perceptions of pupil attributes. In public response situations it may represent an active attempt to get the children to participate and to compensate for their lack of confidence. In private contacts it is in line with the perception of children in the low expectation group as lacking perseverance and task concentration, reacting poorly to new work and needing teacher guidance.

Expectations and Achievement

The teachers did not have the NZCER Progressive Achievement Tests (Reading Comprehension, Reading Vocabulary, and Listening Comprehension) results available when the teacher rankings were collected. However, teachers' expectations for their pupils based on these rankings were strongly related to results on the PATs and this provides evidence that the teachers' expectations were relatively accurate.

Further, expectations were also strongly related to an end of the year composite measure of school achievement based on teacher ratings. Does this mean that a self-fulfilling prophecy was operating? Not necessarily, unless there is clear evidence that those expectations were consistently communicated to pupils in a way that shaped their behaviour and achievement. Here no clear-cut evidence of this nature is available.

What of ethnic group differences? The teachers perceived Maori and Pacific Island pupils in less favourable terms than Pakeha pupils and had lower expectations for their general ability to do well at school. Nevertheless, there were no statistically significant differences in teacher-pupil interaction along ethnic group lines. Thus the teachers expected less of the Maoris and Pacific Islanders than the law of averages suggests they should have, but they treated them similarly to others in their expectation groups.

As an aside we can ask, does this similar treatment mean that the Polynesian and Pakeha children achieved at similar levels? Unfortunately, no. Overall, Pakeha children scored higher on the PAT tests and their end of year school achievement was rated higher by their teachers. While the greatest differences in levels of achievement lay between expectation groups, within these groups Polynesian pupils tended to score slightly lower. This difference may be due to some dimension(s) of behaviour tapped by the Polynesian/Pakeha dichotomy used in this study but not reflected in the teacher expectation measure used. Cultural differences in communication, motivation, and self concept, or preferences for different kinds of learning environments could be involved.

Implications for Teachers

The study described here was essentially exploratory. It was naturalistic rather than experimental and does not encourage definite casual inferences. However, it does provide some important data on the nature of teacher expectations in New Zealand ethnically mixed classrooms and the relationships between teacher perceptions, expectations, and interactions with pupils.

It is encouraging to find no evidence that teachers treat children for whom they have low expectations in a generally negative way. Despite the more negative and stereotyped perceptions of Maori and Pacific Island pupils as a group, teachers seemed to put into practice a philosophy of equitable and fair treatment of all pupils.

(It would be a pity, however, if in the effort to provide fair and equal treatment that teachers are unable to draw upon and react to positive aspects of ethnic group differences. A greater similarity in level of interests, attempts to learn, and a positive self-regard in all children might result.)

Being aware of the possible effects of their expectations can help teachers reduce the likelihood of undesirable self-fulfilling prophecies occurring. It is not appropriate, however, for teachers to try to avoid forming expectations altogether. Even if teachers do not look at achievement records or test information and they avoid discussing pupils with previous teachers they still tend to form strong and general impressions quickly, through interaction in the classroom.

It is also sometimes suggested that teachers should only have positive expectations. This is unrealistic when carried to the extreme — wishful thinking will not get rid of all differences in pupil abilities and interests. More important than having high expectations all the time is that expectations should be appropriate and followed up with appropriate behaviour. There is nothing wrong with expecting that a student will have difficulty if help is provided to remedy the difficulty. Such help should be positive, providing the pupil with courses of action, and the way it is given should avoid communicating negative expectations.

Teachers should continually check their expectations for pupils and adjust them where necessary. It is not always easy. When expectations become strong and stable there is a tendency to notice only those aspects of a person's behaviour that 'fit' the expectation. In addition, the way we interpret what we see may be affected. For instance the behaviour of two children might be exactly the same but the teacher perceives the 'bright' one as concentrating on working out the problem and the 'slow' one as being hopelessly lost.

When teaching children from a variety of ethnic and cultural backgrounds it should be remembered that cultures are systems of shared understandings, especially about what words and actions mean and how important values should be expressed. Unintentional miscommunication and misunderstanding may occur in the classroom if teachers do not know the important differences that exist.

Where does the information we use to judge the likely success or failure of our pupils come from? Often from observation of academic skills and work habits in the classroom. Consider what will happen where stereotyped judgements are adopted on limited information with little attention being paid to how a child is actually behaving. Focussing on a label may switch attention away from teaching and fostering progress. When poor academic skills can be 'explained' by saying that a child is a slow learner or is a member of a particular ethnic group then there is little room for the teacher. What effect can he or she have?

Conclusion

What does this study have to say to New Zealand teachers, particularly those in multi-cultural classrooms? While it is not representative of all those teachers, we think it does provide some indications of what is going on and what to look out for.

Some of the study's findings are heartening, others less so. It is unfortunate that teachers still hold negative stereotypes about Maori and Pacific Island pupils and are more likely to expect them to be of lower ability and to do less well in school than Pakeha pupils. Further, despite the fact that Pakeha and Polynesian pupils have similar kinds of classroom interaction with their teachers they do not achieve at the same level.

On the positive side it is encouraging to find that overt discrimination against minority group children does not seem to be the general case. It is also good to see that teachers are attempting to even up the number of interaction opportunities available to pupils. This is not to say that all children experience the same classroom environment, for the type and quality of interaction is also very important, but it does seem to indicate that the teachers are trying to provide for all their pupils.

What are the things to watch out for? Clearly those who teach children from a variety of ethnic and cultural backgrounds should try to avoid making premature, inflexible judgements on little information. Judging children on the basis of a single characteristic such as ethnic group will result in many inappropriate expectations being formed. It is all too easy to teach on the basis of untested assumptions about children. Check them out and remain open to noticing what your pupils are really doing. Otherwise you may only see what you expect to see.

Notes:

Impressions of Other People
Discussions can be found in:
Downey, M. *Interpersonal judgements in education.* London:Harper and Row, 1977.
Hargreaves, D.H. *Interpersonal relations in education.* London: Routledge and Kegan Paul, 1972.

Teachers and Expectations
Discussions and relevant research are contained in:
Brown, C. Teacher expectations: Sociopsychological dynamics. *Review of Educational Research,* 1976, 46, 185-213.
Brophy, J.E., and Good, T.L. *Teacher-student relationships: Causes and consequences.* New York: Holt, Rinehart and Winston, 1974.
Dusek, J.B. Do teachers bias children's learning? *Review of Educational Research,* 1975, 45, 661-684.
Good, T.L., and Brophy, J.E. *Looking in classrooms* (2nd ed.). New York: Harper and Row, 1978.
Nash, R. *Teacher expectations and pupil learning.* London: Routledge and Kegan Paul, 1976.

Expectations as Self-Fulfilling Prophecies
The controversial study by Rosenthal and Jacobson is in:
Rosenthal, R., and Jacobson, L. *Pygmalion in the classroom.* New York: Holt, Rinehart and Winston, 1968.

A New Zealand Study
St. George, A. *Perceptions, expectations and interactions: A study of teachers and pupils in five ethnically mixed primary classrooms.* D. Phil. thesis, University of Waikato, 1978.

Implications for Teachers
Further helpful suggestions are in Downey, Good and Brophy mentioned above.

Discussions of problems of cross-cultural communication with illustrative examples relevant to teacher expectations and the school setting are found in:
McDonald, G. Working with Pacific Island and Maori parents. *Set,* 1979, No. 1, Item 4.

Metge, J., and Kinloch, P. *Talking past each other: Problems of cross-cultural communication.* Wellington: Victoria University Press/Price Milburn, 1978.