A state of the art survey provided an overview of approximately ten years of applied behavior analysis research in New Zealand. Research areas identified in an earlier review are updated in this report, and current directions in research activities are identified. General programs, including innovative special education facilities and remedial programs, are reviewed. Explorations made of class management and instruction in normal and special education classrooms examined the effectiveness and extent of behavior modification practices used with differing target populations and settings. Issues and concepts in diverse areas, such as behavior change variables, complex behavior and curriculum programing, and professional training and methodology, are examined. A bibliography of reference reading materials on this topic is appended. (JD)
Behaviour Analysis in Educational Settings in New Zealand: Current Research Trends

Ted Glynn
Stuart McNaughton
Department of Education
University of Auckland

State of the art papers can be an important exercise for behaviour analysts. They provide a special opportunity to undertake a fundamental activity, the observation and analysis of repeated measures. With repeated research activities and their outcomes considered as samples of data, a rewarding level of induction is possible. For example, results can be obtained such as establishing the generality of concepts, and identification of potentially significant areas for future research.

This paper is divided into three sections. Firstly research areas identified in an earlier review (Glynn, 1976) are updated, with those studies which add significant information to the data base and raise important issues receiving close attention. A second section outlines current directions in research activities which establish new concepts (thereby, in some instances, defining the limits of previous ones). Finally, more general programmes, including innovative special education facilities and remedial programmes are reviewed.

We have adopted Baer, Wolf and Risley's (1968) classic statement on the dimensions of applied behaviour analysis as a basis for selecting research to review. The major concern is not to make a
critical methodological evaluation of New Zealand research. Rather, we re-
view the procedures and concepts used in, and the outcomes of, available
published and unpublished research reports which have adopted a behaviour
analytic framework. In so doing methodological issues are noted and where
appropriate commented on.

The paper is unashamedly parochial. This position is justified on two
counts. There is enough quality research in New Zealand to justify limiting
the review to local studies. In addition there are characteristics of the
New Zealand educational scene which mean local research may have greater re-
levance (e.g., approaches to teaching reading).

I Replication and Extension of Previous Research
In an earlier paper five years of applied behavioural research in New Zealand
schools were reviewed (Glynn, 1976). Studies were grouped into research on
teacher management, peer management and self management. The majority of
these studies were concerned with the management of problem behaviour in nor-
mal primary and secondary school classrooms using direct intervention pro-
cedures.

Research has continued to be productive in these areas. Many of the
studies come from special projects and theses required for graduate and pro-
fessional training courses. The work of Church at the University of
Canterbury must be recognised as a significant contribution, not only in
training but in editing and publishing through the Educational Research
Newsletter. Studies of classroom contracts, teacher attention to on and
off-task behaviour, self management and other procedures are contained in
the Newsletter.

Research has also extended the established principles and procedures
into different subject and behaviour management populations, across a greater
range of educational settings, and to different behavioural forms. These
are important activities when collectively seen in terms of the tactics of
scientific research described by Sidman (1960) as establishing the reliab-
ility and generality of the data.

Teacher management
Functional relationships between levels of inappropriate classroom behaviour
and amount of teacher attention to appropriate behaviour were demonstrated
in early studies (e.g., Glynn and Quinnell, 1971; Glynn, 1972). It is
therefore important to note recent descriptive data on 'natural' rates of teacher verbal attention to on and off-task behaviour have been published. These were abstracted from baseline observations of 10 teachers in three South Auckland Intermediate schools (Thomas, Presland, Grant and Glynn, 1978). The data show rates of attention were higher to off-task behaviour than on-task behaviour for 9 of the teachers. (Overall rate means were; attention to on-task = 0.20 per minute; attention to off-task = 0.58 per minute. The overall mean of intervals scored on-task was 66 percent.) These rates are similar to those obtained by White (1975) for similar (Grade 7) classrooms in America.

The overall finding of low rates of attention to appropriate behaviour in situations where there is a relatively high rate of problem behaviour has been recently extended to academic tasks. In a pilot study McNaughton, Glynn and Robinson (1980) describe five teachers' interactions with low progress 8 to 12 year old readers during one-to-one oral reading. Over ten sessions the average number of approval comments directed to correct reading responses was four per session compared with close to 75 percent of all errors being attended to (a majority of them immediately).

As Thomas et al. (1978) suggest there is a need for research to identify sources of control over teacher behaviour (e.g., disapproval comments). For example, it is possible that the reinforcement available in short term behaviour change (e.g., corrections of an error following immediate teacher help), even though having possible inappropriate long term outcomes is very powerful. This may be especially so for children with long histories of difficulties and/or making slow progress (c.f. Patterson's 1977 'coercion' hypothesis, and Wahler's 1976 'negative trap' model).

The procedures developed in earlier research on teacher management of classroom behaviour have been extended to other settings and managers. Litter monitors have been trained to remove litter in an intermediate school playground under a group contingency, where a whole class gained access to free time depending on how much litter remained (Presland, 1978). Johnson (1978) successfully trained part-time and full-time staff at a preschool and special care centre for intellectually handicapped children to effectively change their patterns of attention to appropriate and inappropriate social and play behaviour.
A more ambitious study by Starkey and Perkins (1978) involved the entire roll of a primary school in a programme to change street crossing behaviour at intersections. Procedures included an extensive task analysis, modelling of appropriate behaviour and in-class instruction in basic concepts, training of peer supervisors and mastery training of all children (N = 370) on intersections. Special badges were used as consequences for reaching a criterion level of performance. Unfortunately the data indicate a minimal change in the children's appropriate street crossing behaviour. The authors claim this was not due to problems in the basic procedures, but due to insufficient programming for maintenance and generalization.

Fry and Thomas (1976), however, did programme for generalization. They implemented a token economy system (with time out) in an adjustment class to increase the on-task behaviour of behaviourally disordered six-year-olds. Following successful treatment the programme was stopped and the children were placed back in regular classes. The data show successful reintegration occurred for at least six of the seven children. The programming for generalization was not extensive, involving withdrawal of the token programme in the adjustment classroom prior to full placement, approximation of praise and attention for appropriate behaviour to regular classroom levels, and gradual introduction to the normal classrooms.

The trend noted in the earlier (1976) review towards greater use of contingencies on accuracy and academic work completed is now firmly established. Moreover, not only teachers are involved. For example, Fry (1977) used parents at home as tutors for low progress 7-to-11-year-old readers. Parents used praise tokens and tangible reinforcers as consequences, and made and sequenced cards for word recognition training. All 30 children made progress, gaining an average of 8 months in reading age (Burt Word Reading Test) for two months of tutoring.

Fry's data agree with other research using parents as home tutors (Robinson, Glynn, McNaughton and Quinn, 1979). She notes that parents from a lower middle class suburban area both showed strong interest and cooperation, and demonstrated ability to tutor their own children. Parents represent a highly motivated group and a potentially significant resource for remedial tutoring.

Two studies (M.A. theses) show why research has moved further toward concerns for variables directly influencing academic performance. In an ex post facto analysis following an unsuccessful attempt to gain control
over off-task behaviour, Wernham (1978) found little (correlational) relationship between low noise levels and on task behaviour or academic output. Glendinning (1978) introduced a token procedure for on-task behaviour and then academic performance in an adjustment class for emotionally disturbed children. The multiple baseline across both lessons and behaviours indicates that contingencies for on-task behaviour had little effect on academic output, whereas contingencies placed directly on academic output produced greater output and increased attending behaviour (c.f. Marholin and Steinman, 1977).

Both studies show that the relationship between attending behaviour and academic behaviour is not simply that high levels of the former are needed as a precondition before changes in the latter occur. Given that learning in classrooms is necessary then following instructions is defensible only to the extent that other learning (including academic learning) is made possible (e.g., Peters, 1973).

If academic gains are not determined by prior change in on-task behaviour in some classrooms, and, if contingencies placed on academic behaviour change on-task behaviour anyway, there is an obvious conclusion. Our research efforts should be towards analyses of variables directly related to acquisition of academic skills. Bringing classroom behaviour under the control of academic materials may be both more defensible and better behavioural practice (e.g., Marholin and Steinman, 1977). Thus in some classrooms off-task behaviour might be better viewed as an adaptation entirely appropriate to a poor programme or curriculum. In which case the concern of the behaviour analyst should be curriculum management.

Peer management
Several studies already reviewed involved peers as behaviour managers. Presland's (1978) procedures for increasing litter removal from an intermediate school playground included the use of high school students as supervisors and data collectors. They carried out these activities as part of a community studies course. Presland reports anecdotaly that having functioned as supervisors the high school students displayed greater awareness of the litter problem. Peer tutoring and management represent an important area for further research in New Zealand given general findings on effective use and mutual educational benefits for
tutor and tutee (e.g., Dineen, Clark and Risley, 1977; Sanders and Glynn, 1977).

While Starkey and Perkins (1978) had not designed effective procedures for producing large scale changes in street crossing behaviour they do suggest that use of peers may provide the basis for more effective procedures. They anticipate using peer tutors to cue and provide feedback for the appropriate behaviours. Thus they see the solution to their problem of programming for maintenance and generalisation as lying with peers.

Self management
In the area of self management Coleman and Blampied (1977) extended the use of standard procedures to a senior special class for nine 9-to-14-year-old retarded boys (IQ range 51 to 76, Wechsler Intelligence Scale). Self-monitoring, self-recording and self-administering check marks exchangeable for back up reinforcers were associated with increased on-task behaviour. While some trends in the data confuse interpretation (e.g., an ascending baseline) the reversal design strongly supports the authors' conclusion that a mild to moderate degree of intellectual handicap is no impediment to the use of standard self management procedures.

Self management procedures successfully produced high levels of on-task behaviour even though an earlier phase of externally determined reinforcement was not used, thus replicating Jackson and Glynn's (1974) findings. As noted in 1976 (Glynn, 1976) there is still a need to examine the variables within classroom and individual interaction histories which might determine the need for externally imposed contingencies prior to a self management procedure.

Another observation from this study concerns the accuracy of self-monitoring and self-recording. During a phase where the monetary value of back-up reinforcers was substantially increased, mean accuracy of self-recording declined (the mean difference between observers' and children's assessments of amount of on-task behaviour was close to 20 percent). The authors then changed the backup contingency from consumption (i.e., possession) of play materials to access to them in free play. The difference in assessment of on-task behaviour by the class and the observer was reduced while high levels of on-task behaviour were maintained. This suggests the selection and programming of back-up consequences may be a critical component for this paradigm.
A feature of the research into teacher management and self management has been the analysis of components of treatment 'packages' or comparison between such 'packages' or variables. Chiew, Parsonson and Priest (1979) analysed the differential effects of components of the 'good behaviour game' on disruptive behaviour of a third form class during English and Mathematics. Rules, rules plus feedback, and rules plus feedback plus consequences were introduced successively in a multiple baseline across lessons. This design, describable as an A - B - B + C - B + C + D design, or sequential components design, has been used in other component analysis research (e.g., Johnson, 1978).

The data indicate that rules alone (B), and rules plus feedback (B + C - marks on the blackboard) exerted little control over disruptive behaviour. As a corollary, with all components introduced successively and in the particular order noted above, disruptive behaviour was dramatically reduced in both lessons.

The problems of sequence or interaction effects in reversal designs (Hersen and Barlow, 1976; Ulman and Sulzer-Azaroff, 1975) are particularly compelling in component analysis research. Hersen and Barlow (1976) claim there is a design requirement that only one variable should be changed when proceeding from one phase to the next and components need to be directly compared with other components in adjacent phases (as in an A - B - A - B - C - B comparison where B is compared directly with A, and C). However, when three or more components need direct comparison in balanced sequence, excessively long and unwieldy single subject designs eventuate. If N = 1 designs are retained then there appear to be at least three alternative strategies. Two of these are either a succession of studies each making only one or two direct comparisons (thereby necessitating comparability across studies), or, groups of subjects run in different combinations within one study (thereby necessitating comparability across subjects; e.g., Sanders, 1978). The third alternative is to adopt a specialised variant of the reversal design where components alternate within phases (see discussion below).
Coleman and Blampied's (1977) self management study used ticks on paper with (B + C), and without (B), back-up reinforcers. Rather than the sequential design of the previous study, this study introduced phases in the pattern A - B - B + C1 + B + C2 - A - B - B + C3 (C1 = consumption reinforcement, C2 = enhanced consumption reinforcement, C3 = access reinforcement). Because no two conditions share adjacent phases (e.g., B - B + C - B) it is difficult to make strong comparisons. Nevertheless the data do suggest the increased effectiveness of the procedure when back-up reinforcers are available and the superiority of an 'access' contingency to a 'consumption' contingency.

These analyses raise an applied issue. In its simplest form it is the perennial debate of the distinctions between basic and applied research. Azrin (1977) has argued for a low priority to be given to analytic studies, claiming applied efforts should be outcome oriented. But the concentration on outcomes and the construction of 'treatment' packages from known and effective procedures has recently been criticized as limiting (Deitz, 1978). Deitz's claim is that the complete transition to a technology is premature, it is presumptuous to assume that the most useful information about behavioural interactions in applied settings has already been gained. This argument can be extended. The technology for practitioners ultimately may be weakened because analytic, investigative research was not carried out and further significant variables identified. The issue does not have to be resolved in an either/or fashion (c.f., Baer, 1978). That applied research can be both analytic and outcome oriented has been adequately demonstrated by several researchers, including for example Risley (Risley, 1977) and Lovitt (Lovitt, 1976).

Another issue arises from direct comparisons between different variables or treatment packages. While such comparative research is often requested, caution is required. To obtain unambiguous results it is usually necessary to have extensive knowledge of the operating characteristics of variables or treatment packages (Sidman, 1960). This means data are needed on their operation under different conditions and for different values of the dimensions involved. Without this information comparisons are uninterpretable because they may not adequately represent the power of either or both 'packages'/variables.
Given this caution it is interesting to note a comparison that Brown, Parsonson, Priest and Glynn (1979) made between two differential reinforcement of low rate schedules (spaced responding DRL and Interval DRL). These schedules were applied to the disruptive classroom behaviour of two 8-year-old boys. Response patterns characteristic of those obtained in laboratory operant studies did not occur. However the multiple baseline did suggest that the schedules controlled the disruptive behaviour.

McNaughton and Delquadri (1978) compared modelling procedures (e.g., 'the word is...') with a prompting procedure (e.g., 'What is the first sound?') for teacher correction of errors during oral reading. The comparison was designed so that a more proficient learning disabled 11-year-old received the two procedures in the opposite order from a less proficient peer. Conclusions concerning differential effects as a function of inter-individual difference are representative statements and need group (actuarial) data. Nevertheless the difference between the two subjects, namely that modelling was more effective for the less proficient and prompting was more effective for the more proficient reader, warrants further research.

A feature of this study was the use of short term reversals between longer phases of the two procedures in order to reduce carry over effects or contrast effects. The design could be represented as


(where A = no oral reading tutoring, B = prompting and C = modelling). The design contravenes Hersen and Barlow's (1976) strict requirement for direct comparisons from adjacent phases, but may represent a useful defensible exception where phases placed in adjacent positions are likely to be interactive.

II Research trends and emergent concepts

Further delineation of concepts, and shifts in research emphases are apparent in the literature. The studies which show these developments exemplify the breadth of behaviour analysis. The methodological and conceptual approach of behaviour analysis emphasizes idiophenomena in applied settings and induction from replications of the patterns found in repeated measures. This produces a particularly sensitive framework for conducting research. It involves a process of scientific development
analogous to the Piagetian concepts of assimilation and accommodation; (see Piaget, 1978 for his description of these processes at work in scientific endeavours).

The development of the framework can be seen in the four areas of (1) research into learning processes; (2) elaboration of concepts of complex behaviour and curriculum programming; (3) research into the training of professional and para-professional educators; and, (4) technological advances in observational and analysis procedures.

Research into processes of behavioural change

Further identification and conceptualization of variables influencing behavioural change in educational settings continues. Contributing to shifts in our understanding is research which has examined (1) ecological variables and setting events, (2) error correction learning, and, (3) generalization.

The need for applied behaviour analysis to actively research ecological variables has been succinctly stated by Risley (e.g., Risley, 1977; Krantz and Risley, 1977). He claims applied problems can be extended to involve behavioural 'ecosystems' (such as whole schools, institutions etc.). Risley is confident that behaviour analysts can discover general principles of the organization of living environments for dependent populations.

Risley's work in preschool settings also demonstrates that manipulation of ecological variables such as seating arrangements, sequencing of activities, and, naturally occurring adult-child interactions (incidental teaching settings) produce behaviour change which is effective and as efficient as standard behaviour modification procedures. But in addition, such manipulations do not require large expenditures of resources (e.g., teacher training) or the extra response cost to teachers that may occur when individual contingency management procedures are used. A particularly powerful example is the incidental teaching setting where significant and generalised oral language gains have occurred.

In New Zealand O'Rourke and Glynn (1978) have studied ecological variables in an intermediate school playground. Extensive data were gathered on the effects of types of playground equipment and adult participation with equipment on appropriate and inappropriate playground behaviour. Provision of equipment and periodic changes in equipment increased and maintained children's appropriate participation. But adult presence and participation with equipment produced high stable rates of participation irrespective of equipment changes and type of
equipment. In conclusion the authors again note the effectiveness of indirect modification of environments. The manipulation of ecological variables reduced undesirable playground behaviour by generating high rates of incompatible (i.e., appropriate) behaviour.

Classroom or instructional setting events are specific examples of effective ecological variables. A setting event:

influences an interactional sequence by altering the strengths and characteristics of the particular stimulus and response functions involved in an interaction

(Bijou and Baer, 1978, 26)

As defined, the concept emphasises the influences exerted by an individual's history of interactions as well as physical, biological and social factors in the general context. Setting events are therefore variables that function to increase the effectiveness of 'naturally' occurring (i.e., currently established) interactions. They also include variables which could influence the probability of a behaviour occurring without directly modifying specific stimulus and response functions. Thus, given appropriate interactions occurring with some minimal frequency, and that a (functional) setting event can be identified, then direct intervention may not be necessary to increase the frequency of a targeted behaviour. This reduces the programming necessary for generalization, maintenance and instructional independence. These needs often arise from direct intervention procedures.

Much of Risley's work in the management of environments (e.g., Krantz and Risley, 1977) is understandable in these terms. The utility of identifying such events is further illustrated in reading instruction research by McNaughton and Glynn (1979), and, McNaughton and Delquadri (1978).

In the former study a simple manipulation of the timing of a tutor's attention to oral reading errors provided opportunities for self correction to occur (by delaying error correction until the end of a sentence). All 6 of the six-year-old normal readers increased frequencies of self correction. The increased opportunity enhanced the power of variables contained within the text to cue behaviours of self correction. Similarly, in the latter study, an extra 10 minutes
of practice alone produced increases in oral reading behaviours for two learning disabled 11-year-olds.

Not all events which might logically be setting events necessarily have that function (e.g. noise level, Wernham, 1978; 'attending behaviour', Glen-\ldots{}denning, 1978). Setting events are seen as acting on stimulus-response interactions. When functional stimuli associated with the setting events are not present, or when behavioural competencies are restricted, the potential setting event has no function. Thus noise level or 'attending' behaviour as candidates for setting events for academic behaviour depend on other antecedent and consequent stimuli such as appropriate instructional material and teacher behaviour.

Further research is needed which identifies strong and generally functional setting events, the conditions which lead to an event having such a function, and the way setting events can be utilized in classroom programmes. Conditions conducive to indirect forms of behavioural change need to be distinguished from settings where direct intervention (e.g., Direct Instruction, Becker and Engelmann, 1978) is necessary.

There may be some settings where direct forms of instruction may interfere in the long term acquisition of proficiency. Self regulation of academic tasks (see discussion below) or incidental learning of oral language (e.g., Risley, 1977) may be two such areas. With some subjects in some settings close specification and training in component skills may reduce the flexibility and generativeness of the learning. Stokes and Baer's (1977) technique of training for generalization termed 'train loosely' provides a parallel. And, as Ryan (1979) notes, there are data to show similar processes at work when behaviour modification is used in the regular classroom. Inappropriate consequences of restriction in the range of behaviour emitted and the undermining of 'intrinsic' reinforcement can occur.

Aspects of curricula could be analysed as setting events. Comments in some research on classroom on-task or academic behaviour describe high or low data points which are at variance with trends. These are often attributable to a new feature of the teacher's programme. Rather than treating these as error variance perhaps these sources of variability should become crucial areas for analysis. There are examples of this sort of research on academic learning in research outside New Zealand (e.g., Whitehurst, Domash and Di-\ldots{}

Another area of research concerns error correction in complex tasks such as oral reading. There has been a tendency in the research literature to consider attention to inappropriate behaviour (via error correction)
as something to be avoided. While teacher social attention to off-task behaviour has often been shown to have reinforcing properties such may not be the case with complex tasks like reading.

Unintended reinforcing outcomes of error correction have been shown on complex tasks (e.g., Sajwaj and Knight, 1971). But there are both compelling conceptual arguments and some data to suggest error correction (a more general term commensurate with other research areas might be negative feedback) may have an important instructional role. Many academic tasks involve learning to selectively attend to several cues. In more complex tasks these cues provide different sorts of information depending on the combinations of cues present. For example, graphic, syntactic and semantic cues in a reading text.

With these tasks error correction may provide information which enables the student to acquire the skills efficiently and economically. More so perhaps than simply concentrating on scheduling component tasks so that errors are reduced to a minimum and positive consequences are the dominant instructional strategy.

Certainly, considerable academic learning can come from error correction even when there are minimal positive consequences (McNaughton and Delquadri, 1978). This area promises to be an important one for classroom instruction on academic tasks. In the wider literature researchers are beginning to analyse parameters of error correction (e.g., Hanson and Eaton, 1978).

Questions of generalization across settings, tasks (responses) and time have been raised in several studies. Defining criteria for generalization is itself an issue (Robinson and Swanton, 1978) but a working description for our purposes here would be, change in behaviour primarily as a function of behavioural change in a different setting, task or at an earlier time. At least two questions are identifiable in the research. What are the interactional processes associated with all three types of generalization, and, what are the features of tasks and repertoires which determine response classes (i.e., responses that covary such that contingencies applied to one, affect the others)?

These questions have considerable educational importance. For example, generalization is a central issue facing research on parent training (Glynn, McNaughton, Robinson and Quinn, 1980); teacher training
(Robinson and Swanton, 1978); academic learning (McNaughton, Glynn, Robinson and Quinn, 1979) and problem solving (Parsonson, 1978).

In a series of studies Parsonson has been analyzing generalization across academic-related skills. In 1978 Parsonson (1978) reported that studies to date on problem solving and other creative behaviour (e.g., writing) represents beginnings of a research enterprise, but much remains to be done. He emphasised generalization across topographically similar and dissimilar tasks as an area needing further research.

In further generalization research moderately retarded adolescents have been trained to discriminate and label the numbers 0 to 99 (Young and Parsonson, 1979). Modelling and token reinforcement procedures used with exemplars produced generalization to untrained numerals. However, there were marked differences across subjects in the number of training sessions necessary and the rapidity of the emergence of generative labelling.

Similarly, Dawson-Wheeler and Parsonson (1979) taught several reading skills to a 10-year-old 'aphasic' child. They used discrimination training with correction procedures involving modelling, prompting and imitation and descriptive praise. Generalization to untrained reading and writing responses was continually probed within a multiple baseline design. Generalization was obtained from trained reading words, phonemes/digraphs, and sentences, to their written form. But generalization within reading responses to untrained blends and untrained words occurred only after the boy was taught a related set of phonemes and digraphs and not when a set of sight words were trained. This result highlights the questions introduced above, in this case about component skills in reading (sight vocabulary) which are generative and the most appropriate training sequence to produce them.

This question was present in the Mangere Home and School Project where parents were trained to tutor their low progress normal 8-to-12-year-old readers in oral reading at home (McNaughton, Glynn, Robinson and Quinn, 1979). Reading behaviours were tutored at home and generalization to independent reading at school was continually measured. Home tutoring emphasised proficiency in reading meaningful texts (rather than, for example, word recognition c.f., Fry, 1977). Proficiency included accuracy and self correction measures.
While considerable progress across book levels was made at home during tutoring only 2 of the 8 children showed similar changes in reading at school. It seems that generalization was limited to children who were receiving some individualized programmes in their classroom and who had shown the most proficiency (especially in maintenance of self correction) at home.

Concepts of complex behaviour and curriculum development

Studies have added further dimensions to the self management procedures established in research reviewed earlier. One area of concern has been the direct training of skills involved in self management (or more comprehensively self regulation) on academic tasks. Previous research has primarily been concerned with the maintenance and increase of simple (i.e., easily discriminable) appropriate behaviours. Procedures for self monitoring and self administering reinforcement have been relatively obvious, with external support such as intermittent cues and cards for tokens being used. Procedures have tended to emphasize motivational functions of self management.

When self regulation of academic tasks is considered it becomes obvious that procedures for motivation are only one (perhaps secondary) concern. Firstly, on all but the simplest of academic tasks self regulation is dependent on more extensive discrimination learning. That is, as tasks become more complex (as in academic learning), distinctions between appropriate and inappropriate behaviour become more relative and determined by several cues. Appropriate behaviour does not stand in an easy reciprocal relationship with inappropriate behaviour nor is it a clearly circumscribed class of readily identified responses. Solving a mathematics problem or observing and correcting an oral reading error show these complexities.

Similarly, simple maintenance of levels of output as a goal is overshadowed by the possibility of self instruction. It is likely that skills can be identified which not only maintain current levels of proficiency under appropriate conditions but also contribute to the acquisition of new skills. Thus the components of self monitoring, and further components of problem solving become critical. Skills for
identifying errors, difficulties or potential problems, strategies for searching and trying out solutions, and skills for recognising best solutions are needed (e.g., Resnick and Ford, 1978).

These difficulties were confronted by Wilson (1978) in attempting to increase independent sentence writing behaviour by eight mildly retarded children in a special class. Within an ongoing token programme, procedures to increase self-selecting and self-locating words were introduced in a sequential components design. Only when all the components of rules and word lists, praise for self-selection and location, and a response cost for teacher-dependent responding were present, did high rates of target self regulatory behaviour occur. The data indicate that externally mediated procedures which included motivational components were necessary to reduce teacher dependency.

A feature of Cameron and Robinson's (1978) treatment of three 7-and-8-year-old hyperactive children involved training self instruction skills (gaining control over motor movement via verbalization about planning and checking) and self management (self monitoring and self reinforcing) for maths problems. Accuracy increased for all subjects as a consequence of the training and there were indications of possible generalization to untrained behaviours of self correction in oral reading.

Both these studies assumed a direct training paradigm for acquisition of self monitoring and problem solving skills. The remedial research which has concentrated on descriptive praise for self corrections in oral reading can also be seen from this perspective (e.g., Glynn, McNaughton, Robinson and Quinn, 1979). However conditions may be able to be identified which support indirect forms of instruction. For example, children making normal reading progress in New Zealand classrooms develop self correction skills and show changes in the topography of these skills over time, without direct instruction (e.g., Clay, 1979). Given the opportunity and meaningful texts at a suitable level of difficulty, self correction behaviours are high probability behaviours for average progress children with no external cueing or consequences (McNaughton and Glynn, 1979). But as suggested previously, direct instruction may be a necessity when incidental learning does not occur.

An implicit question in much of the discussion thus far concerns how complex behaviours such as oral language and academic skills are
conceptualized; both as fully acquired response systems and during acquisition. The issues within this formidable question revolve around the relationship between logical task analyses and empirically determined acquisition sequences under various instructional milieux. A dominant influence in behavioural research has been models of complex skill learning which concentrate on the successive and orderly acquisition of units of skills (e.g., Becker and Engelmann's (1978) analysis of language and reading). Other models which are entirely amenable to a behavioral analysis are possible, including concurrent and mutually interactive learning of several skills. This appears to be a characteristic of how New Zealand children learn to read (e.g., Clay, 1979). (It is interesting to note Panyan and Hall, 1978, recently produced favourable comparisons of concurrent skill learning with serial sequencing of academic tasks.)

Continuing this line or argument further, there is a case to be made for descriptive research of the types of learning possible, and strategies different children use in learning different tasks under different instructional systems. Resnick and Ford (1978) term this research empirical analyses of specific tasks, and argue that it adds considerable information, qualifies logical task analyses, and enables the construction of more effective instruction. Risley (e.g. Risley, 1977) has made a similar argument for developing ecological procedures from descriptions of 'naturally' occurring adult-child interactions.

As a final point, we reiterate comments made in the earlier (1976) review:

The technology of behaviour analysis has a special contribution to make in the field of curriculum development and evaluation... It is important that persons with skills in the analysis and modification of human behaviour and those with skills in the development of curriculum methods and materials should begin to collaborate instead of working in splendid isolation.

(Glynn, 1976, p.92 ff.)

Behaviour analysis has both a technological and analytic contribution to make to research in this area, as well as a data base on teacher behaviour management for the construction of effective curricula.
Research into the training of professional and paraprofessional educators

Procedures for training teachers, parents and other resource personnel in
behaviour modification and behaviour analysis techniques have been examined
in a number of studies. In a standard sequential components design
Johnson (1978) compared theoretical lessons, theoretical lessons plus
practice in analysis and contingency management, and theoretical lessons
plus practice plus feedback about staff and child behaviour. The greatest
change in teacher behaviour occurred with the combination of all three
components (introduced in the particular sequence described above).

Robinson and Swanton (1978) note that there are few follow-up studies
on the generalization of teacher training that fulfil basic methodological
requirements (e.g., several data points and reliability measures for the
follow up data). In a follow-up one to two years after training in class-
room management procedures Robinson, Swanton and Raethel (1979) observed
that two of four intermediate school teachers showed generalization across
students, lessons and time. Clear differences between the two groups of
teachers' attitudes were revealed by interview and questionnaire data. The
authors tentatively conclude that selected attitudinal variables may dis-
criminate between teachers who achieve generalized behaviour change and
those who do not. If this finding holds up with replication then there
are important implications, both for the assessment of generalization and
for the training of teachers.

Procedures for producing short term generalization were used in the
Mangere Home and School Project in training parents in the use of tutoring
techniques for oral reading with low progress 8-to-12-year-olds (Glynn,
McNaughton, Robinson and Quinn, 1980). Instructions were provided in
the form of 'should' statements (e.g., 'I should pause before correcting an
error'). Tutors were trained individually by trainers in a tutoring
session. The trainers cued recall of statements before tutoring, recall
of tutoring performance after tutoring had occurred, and, provided feedback
on the match between statements and performance. With additional early
modelling of techniques and weekly feedback sessions parents learned the
tutoring skills and successfully produced them on occasions when trainers
were not present.
Technological advances
Many of the issues which have contributed to advances in designing research have been outlined in previous sections. The two major areas of development are in designs which handle measures of generalization, and designs which allow for analyses and comparisons of components of 'packages', curricula etc.

In the former case multiple baseline designs with different types of probe measures are emerging as useful designs (e.g., Dawson-Wheeler and Parsonson, 1979; Parsonson and Baer, 1979). Recently a systematic description and evaluation of one variant, the multiple probe-technique has been published (Horner and Baer, 1978). These designs are needed, given shifts towards analyses of more complex behaviour where response class boundaries and precurrent behaviours need to be identified.

Possible approaches in the second area, that of component analysis, have been discussed earlier. A group of designs which seem particularly useful for classroom settings are those variously labelled, multi-element baseline design (Ulman and Sulzer-Azaroff, 1975), multiple schedule and concurrent schedule design (Hersen and Barlow, 1976), simultaneous treatment design (Kazdin, 1978), and, alternating conditions design (Church, 1976). The common element in these designs is the systematic and counterbalanced variation of procedures or stimulus conditions within a phase. The differences between these designs are basically in the timing of the alternation of conditions (e.g., within or across days) and the patterning and consistency over time of stimulus conditions.

Observation and analysis procedures are also showing further development. The complexities of defining and observing play behaviour in preschool children have been discussed by Ballard (1979). The need to record the content of verbal interactions in some settings will necessitate the use of procedures for taking verbatim samples, continuous recording, or use of sophisticated recording apparatus such as transmitter microphones and time set tape recorders. The whole area of analysis of interactional data poses considerable problems for the behaviour analyst although there are some beginnings (e.g., Gewirtz and Boyd, 1976; Patterson, 1977).
The analysis of time series data generally, and the use of statistical procedures, is a topic of considerable debate (e.g., Jones, Vaught and Weinrott, 1977). In a local unpublished paper Arvidson (1977) examined approaches to time series analysis, specifically discussing the issues of serial dependency and other problems facing researchers using ANOVA techniques. He suggests where appropriate the use of regression lines to determine trend and, for statistical comparisons with t tests between phases, using a mean projected from the regression line rather than the obtained mean. These suggestions have been used in two theses (Cameron, 1978; McNaughton, 1978).

III General Research and Treatment Programmes

There are several examples of larger scale intervention programmes which have specifically been designed with research as well as treatment components. The feature of direct and continuing participation in educational settings produces problems of guarding both research and treatment/service responsibilities, and vulnerability to variables that are outside the control of the researchers. Because of this they are valuable sources of information on processes in applied settings. Such programmes should generate more useful behavioural data than the piecemeal approach which has often been restricted to a few university based researchers (Church, 1975). This possibility arises out of the continuing dialogue between research and treatment issues within the setting.

The Mangere Guidance Unit (Thomas and Glynn, 1976) was set up as a systematic integrated behaviour analysis approach to handling behaviour and learning problems. It operates with cooperation between trained teachers, Education Department Psychologists, University Researchers and three Intermediate schools in Mangere. Children are referred by Psychological Services. The Unit carries out treatments in classrooms or settings where the problem behaviour occurs, carries out teacher training, and involves parents in treatment procedures where possible.

The Unit has produced several research reports (Thomas, Pohl, Presland and Glynn, 1977; Thomas, Presland, Grant and Glynn, 1978; Glynn, Thomas and Wotherspoon, 1978), as well as a comprehensive evaluation report to the Director-General of Education (Thomas and Glynn, 1976). The procedures which have been developed at the Unit represent an important ongoing resource.
For example, other psychological services offices and schools have moved towards the concept of a guidance or learning unit (e.g., Glen Eden Psychological Services Learning Unit).

However, experiences with the Unit have highlighted the difficulties in implementing behavioural technology at a school level, rather than at an individual class level. As one of us wrote in 1978:

There remains a large range of political, administrative and communication problems to be solved before behavioural research in New Zealand could claim to be successful at the school level. These problems include the maintenance of regular commitment by Principals and Administrators, (rather than simply initial approval of a research project), coping with the continued and unpredicted staff changes including principals and deputy principals throughout a longer term research, with the resulting vacillation in support for the project. If we are to meet Risley's challenge for behaviour intervention in wider ecological systems (Risley, 1977) then clearly two approaches are possible. The first involves intervention in settings where political support is optimal, so that implementation of behavioural procedures is under the direct control of the researchers. The second approach, that adopted by the Mangere Guidance Unit, involves intervention in settings where political support is more difficult, where implementation of behavioural procedures requires successful enlisting of support from principals and senior staff, hopefully as a result of demonstrated effectiveness of some procedures. The second approach, while potentially more convincing to other teachers and administrators presents a major challenge to behavioural researchers. We have found this situation to be severely punishing for psychologists, subjected to frequent criticism and from senior and junior staff not directly involved with specific programmes in a school. This situation may be overcome only by a heavy commitment to public relations and communication with a school, through participation in regular staff meetings, and being available in the school for individual staff contact. Access to such channels of communication demand the continued goodwill and support of principals and senior staff – difficult in situations where staff turnover is rapid.

My belief is that in the long run there may be a better return from adopting the first approach – that of the behavioural researcher demonstrating what can be achieved in a school setting with optimum political support, with the political support specified as a necessary component of the intervention procedure.

(Glynn, 1978, pp.161-162)

Another, more exploratory project arose out of recommendations in Thomas and Glynn's (1976) evaluation of the Mangere Guidance Unit for greater use of parents. The Mangere Home and School Project was set up to investigate whether parents of low progress 8-to-12-year-old
readers could be trained to successfully tutor their children at home.

It was a project funded as a piece of research for a limited time period. Nevertheless, it has lead to a series of training films, shown on National Television, which have become publicly available as an edited training film (NZCER, 1979) and an accompanying instructional booklet (Glynn, McNaughton, Robinson and Quinn, 1979).

These dissemination outcomes are based on behavioural analysis research carried out during 1978. A research monograph is currently in preparation following the analyses of follow-up measures taken after a year. Several working papers have been presented in the intervening period (Glynn, McNaughton, Robinson and Quinn, 1980; McNaughton, Glynn, Robinson and Quinn, 1979) plus a brief outcome article oriented to teaching personnel (Robinson, Glynn, McNaughton and Quinn, 1979).

The project has produced important data on training parents, effectiveness of parent tutoring at home, generalization of academic behaviour gains across settings, and, the design and refinement of procedures for both tutoring and the observation and analysis of tutored reading.

The purpose of identifying and describing the project is to (somewhat personally) illustrate a claim made by Risley (1978) that significant research is no longer done by individuals and collaboration is needed if the full extent of an applied problem is going to be tackled. There are three dimensions to this need for team research. Applied problems are complex and unwieldy needing considerable input of research expertise over time. Full solution or at least movement towards firm conclusions demand analysis at several levels. For example Clark, Green, Macrae, McNees, Davis and Risley (1977) describe a research programme starting with descriptive data on the extent of problem and degree of interest in problem solution, through programme and component analyses, to marketing research to determine and evaluate the dissemination procedure, resultant effectiveness and consumer satisfaction. A less comprehensive but similar approach was made possible by the team of four researchers working on the Mangere Home and School Project.
Applied problems also make multifaceted conceptual and methodological demands on researchers. The cross fertilization, and testing and revision of approaches available in a team approach proved invaluable in conducting the parent tutoring research.

Finally, a team approach to research increases the probability of an ongoing generative programme of research being produced. The comment that Church (1975) has made, which was reiterated by Risley (1976), is that research has tended to be piecemeal and isolated. The identification and definition of issues, extension and elaboration of concepts, and improvement and advances in methodology are very difficult to achieve outside of an integrated programme of research. Again, the Mangere Home and School Project is a very limited example of the alternative, arising as it did out of previous behavioural research. It has provided a basis and impetus for further research such that each member of that team is conducting further research as a consequence of experience and data from the project.

Summary and Concluding Comments

The cumulative effects of close to ten years of applied behaviour analysis research on educational practice is impressive. Contributions have been made to classroom management, and instruction in normal and special education settings across a large range of ages and skills. The effectiveness of procedures continues to be researched and extended to different behaviour modifiers, different target populations, different behaviours and different settings.

Thus the title of this paper which refers to educational settings reflects part of the growth and demonstrated significance of applied behaviour analysis. But the framework of behaviour analysis is also advancing. Issues and concepts in diverse areas such as behaviour change variables, complex behaviour and curriculum programming, professional training and methodology show the field continues to grow and undergo changes.

This review has only covered a small area within the purview of behaviour analysis and has been selective and restrictive in the research reviewed. For example, research in the Psychology Departments at the University of Canterbury and University of Auckland is producing information on a range of special education training procedures (e.g., oral...
language training with retarded children, Leo, 1978; Miller, 1977; Williams, 1978). Additionally there are important implications in much of the research reviewed which demand further discussion (e.g., issues of graduate training and research funding).

One major area is that of the setting up of programmes of research. This paper commenced with optimism, excited by the possibilities of making connections across studies which might produce grand generalisations. While this serves a useful purpose there is an obvious need for integrated programmes of research which make continuing and significant contributions to the field.
References

Arvidson, G.L. (1977) *Statistical Considerations in Behaviour Modification*. Auckland: Education Department, University of Auckland. (Unpublished manuscript)


McNaughton, S. and Glynn, T. (1979) Self Regulating Processes in Early Oral Reading: Controlling Accuracy via Self Correction. Auckland: Education Department, University of Auckland. (Unpublished manuscript)


Thomas, J.D., Presland, I.V., Grant, D. and Glynn, T. (1978) Natural Rates of Teacher Approval and Disapproval in Grade 7 and Grade 8 Classrooms. *Journal of Applied Behaviour Analysis, 11,* 1, 91-94.


