10/85

Randomised Clinical Trial of Brief Behaviourally-oriented Interventions for Pre-delinquents

Research Project No. 10/85

Robert Sanson-Fisher, Ph.D., and Selina Redman, Ph.D.

Faculty of Medicine, University of Newcastle

Study 1. Abstract

The aim of this study was to design and standardise a brief teacher rated behaviour scale suitable for the assessment of conduct disorders in Australian pupils aged 10 to 15 years. This scale consisted of 35 items which were completed by selecting a response from a 5-point Likert scale ranging from never to very often. Standardisation took place in two stages: reliability and validity testing. Reliability comprised inter-rater and retest agreement and internal consistency. Validity testing involved testing for construct validity. Inter-rater and test-retest reliability was estimated using Cohen's kappa. Nine items with either non-significant ($p>.005$) inter-rater or test-retest kappa statistics were considered not reliable and were excluded from further analyses. The 26 significant items were subjected to a principal components factor analysis with a varimax rotation. This method identified five factors with eigenvalues greater than 1. These factors accounted for 40%, 13%, 7%, 5% and 4% of the total variance. They were named on the basis of manifest item content as acting out, distractability, peer relations, immaturity/withdrawal, and delinquency respectively. Cronbach's alpha was conducted on these factor based subscales. Alpha coefficients ranged from .47 to .92. It is concluded that the 26-item scale has several advantages. A major advantage is its suitability for an Australian sample. It is recommended that the scale be used in conjunction with broader based scales such as the Achenbach Teacher Report Form if discrimination of behaviour problems is required.

Study 1. Development and Standardisation of a Teacher Rated Behaviour Scale for the Assessment of Conduct Disorders

INTRODUCTION

The aim of this study was to design and standardise a brief teacher rated behaviour scale suitable for the assessment of conduct disorders in Australian pupils aged 10-15 years. Conduct disorders refer to behaviour which persistently violates the rights of others and fails to meet age appropriate societal norms and rules (1). Conduct disorders are an important research area due to their salience, the concern they arouse and the seriousness of their counterparts in adulthood (2). As such, they have warranted more long term outcome research than have most other childhood behaviour problems (2). While other scales exist for which data on reliability and validity are available (3), they do not focus on early detection. Because of the current emphasis on prevention (4,5,6), the present scale was designed for early detection of predelinquent conduct disorders. To determine its usefulness as a research tool, standardisation of the scale concentrated on demonstrating reliability and validity.

METHOD

Instrument Development

The initial constructs were generated from three sources: (1) review of the theoretical literature pertaining to classification of childhood behaviour disorders and, in particular, conduct disorders; (2) review of existing scales and (3) unstructured interviews with school counsellors. This method produced six broad constructs including peer relations, immaturity, compliance, withdrawal and acting out.

A pool of 367 items was generated by review of six existing scales. These items were then grouped according to the construct they were purported to be measuring. On the basis of a set of criteria, the item pool was reduced from 367 to 70 items. A 5-point Likert response scale, ranging from `never' to `very often' was constructed for each of the items. The questionnaire was reviewed on psychologists from the Discipline of Behavioural Science at the University of Newcastle. The 70 item questionnaire was then reduced to 35 items on the basis of responses from 19 special education teachers. Teachers were asked to reduce the item pool by half on the basis of three criteria, including representativeness, sensitivity and acceptability. Finally, the questionnaire was again reviewed in the Discipline of Behavioural Science before distribution to teachers.

Standardisation

Standardisation of the behaviour rating scale took place in two stages: reliability and validity testing. Reliability included inter-rater and test-retest agreement and internal consistency. Validity involved testing for construct validity of the scale.

Reliability Study

Study Setting and Design

The data were collected during November and December, 1985. A total of 8 high schools in the Newcastle N.S.W. region were approached to participate in the study. These schools were chosen because they were referring students to the Endeavour Centre intervention programme being conducted by the N.S.W. Department of Education.

Sample

Questionnaires were completed by Maths and English teachers of years 7 and 8. These teachers were selected because they have maximal contact with pupils. Teachers were recruited by the senior English teacher of each school. Questionnaires were completed on randomly selected pupils.

Procedure

Inter-Rater Reliability

Five English teachers per school were asked to complete between 4 to 6 questionnaires each. Questionnaires were distributed with a covering letter explaining the purpose of the study and the randomisation procedure. Teachers were asked to complete the questionnaires on the first and last two pupils on an alphabetical roll. Questionnaires were then distributed to the pupils' Maths teachers. Therefore, separate ratings by English and Maths teachers for the same pupils were obtained, allowing inter-rater reliability to be assessed.

Test-Retest Reliability

Two schools were randomly selected to complete retest measures. Time 2 measures were distributed at these schools nine days after the original distribution of questionnaires. English teachers were asked to again complete the questionnaires on the same pupils. English teachers only were used to control for possible variations due to school subject. Therefore, ratings by the same person on separate occasions were obtained, allowing test-retest reliability to be assessed.

Validity and Internal Consistency Study

Sample

Questionnaires were obtained on a total of 179 clinically referred children. Questionnaires were obtained on referred children only to reduce variability. They were collected as part of the intake procedures at the Endeavour Centre between June 1985 and December 1985. Questionnaires on Primary School pupils were completed by class teachers. Questionnaires on High School pupils were completed by a teacher deemed by the school to have maximum contact with the pupil at that time.

RESULTS AND DISCUSSION

Reliability of Items

Sample

Of the eight schools approached, seven consented to participate, a consent rate of 88%. At time 1, 150 questionnaires were distributed to English teachers and 150 to Maths teachers. 117 questionnaires were returned from English teachers (response rate - 78%) and 116 questionnnaires were returned from Maths teachers (response rate - 77%). Of these, 107 pupils had separate ratings from 27 English and 26 Maths teachers. Thus, each teacher completed an average of four questionnaires.

Both schools approached for test-retest measures, consented to participate. Of the 50 questionnaires distributed, a total of 42 were returned, a response rate of 84%. Of these, 41 pupils had separate ratings from nine English teachers at time 1 and time 2. Thus, each teacher completed an average of five questionnaires at both times.

Analyses

Inter-rater and test-retest reliability of the 35 items was estimated using Cohen's kappa (10). This analysis has been recommended for use as a method of assessment of inter-rater and test-retest reliability (7,8,9). Weighted kappa was used because it is exactly equivalent to the intraclass correlation (8).

Inter-rater Reliability

Kappa statistics ranged from 0.18 to 0.45. The kappa values for nine of the items were not significantly different from 0 (at $p > .005$) indicating that agreement between observers was not any better than that that would be expected by chance.

Test-retest Reliability

The average length of time between test and retest responses was 5.9 days. Kappa statistics ranged from 0.24 to 0.91. The kappa value for one item was not significantly different from 0 (at $p > .005$) indicating that agreement

between observers on this item was not any better than that that would have been expected by chance.

Thus, the nine items with either non-significant inter-rater or test-retest kappa statistics were excluded from further analyses. Notwithstanding the low kappa values, the remaining 26 items with both significant inter-rater and test-retest kappa ($p < .005$) were considered reliable enough to use in future analyses.

2. Construct Validity and Internal Consistency

Sample

This sample consisted of 179 clinically referred children. The sample was composed mainly of boys aged 10 to 13 years.

Construct Validity

The 26 items with significant inter-rater ($p < .005$) and retest reliability were subjected to a principal components factor analysis and varimax rotation using the BMDP statistical package (P4M) (11). This procedure allows the delineation of the nature and number of dimensions which underlie respondents' views of child behaviour (12). 165 cases had complete data. Thus, the response to variable ratio was 6.4:1, a ratio deemed acceptable in previous research (13,14). This method identified five factors with eigenvalues greater than 1. These five factors accounted for 40%, 13%, 7%, 5%, 4%, respectively, of the total variance. Thus, these five factors accounted for 69% of the total variance, a proportion which has been deemed acceptable in previous research (15).

The retained factor based subscales were named on the basis of manifest item content. These are shown in the following tables of sorted rotated factor loadings.

Subscale 1 - Acting Out

The first factor contained items which related to overt behaviour. This factor which accounted for 40% of the variance indicates that these types of behaviours are most salient for teachers. Other factor analytic studies indicate this type of factor as the major subscale (16,17). On the Teacher Report Form, the Aggressive subscale has the highest loading on the externalizing broad band scale. The Acting Out subscale, is similar to the Conduct Problem factor in the revised Teacher Rating Scale and the Aggressive factor in the Teacher Report Form. These factors have a common core of items reflecting arguing, temper tantrums and uncooperative and oppositional behaviours (18). The Acting Out subscale however, does not encompass sulky moody behaviours like the other two scales. In addition, the Acting Out subscale resembles the conduct disorder scale in the revised Behaviour Problem Checklist (17).

TABLE 1

Subscale	No.	Item	Factor Loading
Acting Out	31	Is impertinent, cheeky	0.85
	5	Interrupts, calls out and inappropriately talks to other children in class	0.80
	4	Is overactive or restless	0.76
	22	Disobeys requests	0.70
	21	Displays unpredictable explosive behaviour	0.67
	34	Disturbs other children by teasing, interrupting, provoking fights	0.65
	19	Complains s/he is discriminated against	0.60
	3	Has temper tantrums	0.60
	8	Associates with others who get into trouble	0.53

Subscale 2 - Distractability

Items which related to concentration and attention made up the second subscale. This factor is similar to the inattentive-passive scale of the revised Teacher Rating Scale (16), the inattentive scale of the Teacher's Report Form (19), and the attention problems-immaturity scale of the revised Behaviour Problem Checklist (17). Unlike these, however, immature, childish behaviours do not load on this scale.

TABLE 2

Subscale	No.	Item	Factor Loading
Distract-ability	2	Displays poor concentration in academic work	0.83
	16	Has short attention span	0.80
	13	Behaves in an apathetic unmotivated manner	0.78
	14	Does not listen carefully to instructions or explanations	0.75
	33	Shows little initiative	0.73
	1	Does not complete assigned tasks	0.58
	30	Is uncooperative in group situations	0.54

Subscale 3 - Peer Relations

Items loading significantly on the third factor were concerned with poor peer relations. This is an important area of child development (20). Poor peer relations have been suggested to be linked to a number of negative outcomes, including school maladjustment and drop out, delinquency, adult mental health difficulties and lower levels of academic performance (21). This subscale compares with the unpopular scale of the Teacher Report Form (22). While the Peer Relations subscale includes aggressive peer behaviour (e.g. fights, bullies), the Unpopular scale includes general items of a less

behavioural nature (feels persecuted, feels worthless, lonely). The Bristol Social Adjustment Guides include a similar scale named peer-maladaptiveness (23).

TABLE 3

Subscale	No.	Item	Factor Loading
Peer Relations	15	Fights with other children	0.79
	25	Becomes involved in quarrelling, squabbling or bickering	0.76
	26	Is teased by other students	0.71
	24	Displays angry behaviour when teased by other students	0.68
	12	Bullies other children	0.65

Subscale 4 - Immaturity/Withdrawal

The fourth factor labelled immaturity/withdrawal related to internalizing or withdrawal behaviours and accounted for only 5% of the variance. This suggests that teachers are less concerned with these types of behaviours than those which are more disruptive to class discipline. This scale compares with the social withdrawal scale of the Teacher's Report Form (24). Unlike the social withdrawal scale, however, it also encompasses immature behaviour.

TABLE 4

Subscale	No.	Item	Factor Loading
Immaturity/ Withdrawal	20	Acts young for his/her age	0.42
	28	Withdraws socially	0.70
	29	Sulks	0.66

Subscale 5 - Delinquency

The last factor labelled Delinquency accounted for only 4% of the variance. Conduct disorders may be of a delinquent or non-delinquent or pre-delinquent type (30). Because of the current orientation towards prevention of conduct disorders (4,5,6), this scale aimed at assessing behaviours indicative of pupils at risk of future problems rather than those already displaying delinquent behaviour. The small amount of variance explained by these items indicates that they are only of minor importance to teachers compared with other behaviours.

TABLE 5

Subscale	No.	Item	Factor Loading
Delinquency	35	Truants from class	0.85
	27	Tells lies	0.56

Compared with the apriori scales, a separate delinquency scale was created. Items reflecting withdrawal and immaturity loaded together instead of separately. In addition, items grouped under the apriori compliance scale loaded onto the acting out scale. This means that teachers see these type of behaviours in the same way rather than separately.

Internal Consistency

For factor based groupings of items to be useful as subscales in a questionnnaire, they must be reliable measures of the underlying dimensions, i.e. they must be internally consistent (26). Cronbach's coefficient alpha has been recommended for use as a method of assessment of internal consistency (27). Minimum alpha coefficients of 0.5 are recommended for subscales which are to be used for group comparisons (27). The Cronbach's alpha coefficients for each subscale are shown in Table 6. These ranged from 0.47 for items

loading on the fifth factor to 0.92 for the first factor-based scale. The alpha coefficient of 0.47 for factor 5 is just below that recommended by Nunnally (27) and Helmstader (28), but the subscale was retained as the questionnaire is in the early stages of development (29). However, if this scale is to be used in future research, further examination of its reliability is recommended. The reliability of the other scales in the present research appears adequate.

TABLE 6

Cronbach's alpha for factor-based subscales

Scale	Label	Alpha
1	Acting Out	0.92
2	Distractability	0.87
3	Peer Relations	0.89
4	Immaturity/Withdrawal	0.67
5	Delinquency	0.47

CONCLUSIONS AND POSSIBLE IMPLICATIONS

It is concluded that the scale has several advantages. It is brief and easy to administer. It has been shown to be suitable for an Australian sample, unlike the majority of existing scales which have not been standardised for Australian pupils. Factor analysis indicates that the scale has construct validity. In addition, Cronbach's Alpha indicates that each factor-based subscale is internally consistent. Because the scale is designed specifically to identify conduct disorders, it will not discriminate between different types of behaviour problems. It is recommended that broader based scales such as the Achenbach Teachers' Report Form be used in conjunction with this scale if discrimination of behaviour problems is required. Further refinement of the scale might examine validation procedures based on external criteria, in particular sensitivity and specificity.

Study 2. Abstract

The aim of this study was to evaluate a short-term prevention orientated cognitive-behavioural intervention which focusses on conduct disordered pupils aged 10 to 15 as its target population. The study comprised a randomised clinical trail with one and six month follow-up. The sample was collected from 22 schools in the Newcastle area, including 10 primary schools and 12 high schools. Data were collected from three sources including parents and guardians; pupils; and teachers. Parent measures included the Achenbach Child Behaviour Checklist. Pupil measures included the PAT Comprehension Test, Progressive Achievement Tests in Mathematics and a modified version of the Achenbach Youth Self Report Form. Teachers completed the Achenbach Teacher's Report Form and the Endeavour Behaviour Rating Scale. While some non-significant improvements are noted, the results support research which fail to indicate global improvements, maintenance and generalisation. Parent measures fail to show a statistically significant improvement in experimental as compared with control group pupils. Similar results are found for pupil measures; however, a non-significant improvement in experimental pupils' maths scores is noted. Teacher measures also do not indicate a statistically significant improvement for experimental pupils. However, a non-significant improvement is noted in the short term in more specific conduct (as measured by the Endeavour Behaviour RAting Scale total behaviour score). These results suggest a number of possible recommendations for future interventions. These include ensuring that selection criteria are met; incorporating environmental support (e.g. from parents and teachers); lengthening the intervention period; and incorporating more intensive programming for generalisation.

Study 2 Evaluation of the Endeavour Centre Programme for Conduct Disordered Pupils

INTRODUCTION

The aim of this study was to evaluate a short-term prevention orientated cognitive-behavioural intervention which focusses on the conduct disordered group as its target population.

A number of issues highlight the need for intervention into childhood behavioural problems and in particular conduct disorders. These include the relative high prevalence of these types of disorders (30), their predictive power upon adult life (31,32), the problem they represent for teachers in the classroom (33) and the current shortage of effective referral agencies (5,6,34). The current trend in dealing with behavioural problems is towards a cognitive behavioural approach (35). Most reviews agree that although some studies have reported promising results that merit further investigation, methodological improvements are necessary before the clinical utility of cognitive behaviour therapy with children can be demonstrated (21,36,37).

METHOD

Design

The study comprised a randomised clinical trial with one and six month follow up. Data were collected over a 16 month period between March 1986 and June 1987.

Sample

The sample was collected from 22 schools in the Newcastle area, including 10 primary and 12 high schools. Schools were recruited via application to or selection by the N.S.W. Department of Education.

A selection committee from each school was asked to select pupils in years 5, 6, 7 and 8. Each committee provided approximately 6 names, 3 of which were randomly allocated to the control group and 3 to the experimental group. This method generated a total list of 141 names.

Selection committees were asked to select pupils of average and above ability who were disruptive in the school environment and had failed to respond to normal school disciplinary measures. Because the Endeavour Centre is prevention orientated and not a crisis centre, selection criteria excluded students who: were experiencing severe family trauma; were in crisis or impending suspension; had delinquent convictions; had a history of institutionalisation; were currently involved in therapy with a health professional. Thus, selection was aimed at at-risk predelinquent rather than delinquent conduct disordered students.

Pretest and Outcome Measures

Data were collected from three sources, including parents and guardians, pupils and teachers. Data were collected at pre-intervention plus one and six month follow-up periods.

(i) Parent Measures - Achenbach Child Behaviour Checklist (38)

The Child Behaviour Checklist (CBCL) is designed to assess in a standardised format the social competence and behaviour problems of children aged 4 to 16 as perceived by their parents or others who know them well (24). Social competence is assessed by up to 40 questions measuring the amount and/or quality of the child's involvement in sports, non-sport activities, organisation, jobs, friendships, family and school. The behaviour problems section lists 118 childhood problems and parents are asked to rate how often each occurs on a 3-point response scale (24). The CBCL has been demonstrated to have satisfactory psychometric properties (38).

(ii) Pupil Measures

(a) Progressive Achievement Tests in Reading (PAT) (39)

The PAT Comprehension Test was administered to measure changes in comprehension skill. This test comprises 97 multiple choice items and two alternative forms (Forms A and B) are provided. It is suitable for years 3 to 9 inclusive. Each year-level completes approximately 40 questions. Validity and reliability statistics are available (39).

(b) Progressive Achievement Tests in Mathematics (PATMATHS) (40)

This series of tests is designed to determine basic skills and understandings of mathematics. It comprises three tests, each with parallel forms. This study uses tests 2A and 2B. The tests are designed for years 3 to 8 inclusive. Items have been extensively reviewed to ensure conformity with current Australian curriculum practices and applicability to the Australian context. The tests were standardised in 1983 to obtain Australian norms. Reliability and validity statistics are available (40).

(c) Pupil's Report Form

This questionnaire is based on a preliminary modification by the researchers of the Achenbach Youth Self Report Form (41). It comprises 30 self-report items designed to measure skills and attitudes important to success in school. It was developed in response to results from a pilot study of 22 clinically referred children in the Newcastle area.

(iii) Teacher Measures

(a) Achenbach Teacher's Report Form (22)

The Teacher's Report Form (TRF) is a behaviour checklist designed to obtain teachers' reports of school performance, adaptive functioning and problem behaviour (19). School performance is rated on a 5-point scale for each academic subject. For adaptive functioning, teachers rate children on a 7-point scale in four areas: how hard the child is working; how approximately s/he is behaving; how much s/he is learning, and how happy s/he is. The TRF

also includes 118 behaviour problem items modelled after the parent version of the CBCL (42). The TRF has demonstrated reliability and validity (22).

(b) Endeavour Behaviour Rating Scale

The development and psychometric properties of this scale are reported in Part 1.

Endeavour Centre Intervention

The programme is conducted in approximately 7 cycles per year, with each cycle being 7 weeks duration. The Centre services an average of 4 schools per cycle with each school referring 3 pupils. The outline of each cycle is discussed below.

Week 1

1.1 Selection and Assessment

Criteria for pupil selection has been discussed previously. Admission to the programme was voluntary. Each consenting pupil contracted to enter the programme. Academic and behavioural assessment was conducted prior to entry to the programme. These measures served as pre-intervention evaluation measures. In addition, they served to stream each pupil into appropriate achievement levels for academic classes.

Week 2 and Week 4

1.2 Programme at Endeavour Centre

Pupils attended the Endeavour Centre for 4 days per week. On Friday, they returned to school to practice skills acquired at the centre. The purpose for this was to enhance generalisation. The programme focussed equally towards academic remediation and personal and social development. The programme's timetable was conducted on a schedule similar to a regular school day, with each session being 30 to 40 minutes long.

Academic remediation concentrated on Maths, English and Social Science. Where possible, classes were streamed according to academic ability to provide a more individualized programme for each pupil. These subjects were

concentrated on because they were regarded as central to success in school.

Week 3

1.3 Camp

The camp was conducted over a 4-day period in the second week at various locations in the Hunter and Lower North Coast Regions. Activities included visits to dairy farms, coal mines and fishing trawlers. The major aim of the camp was to promote cooperative group behaviour. In addition, social science skills were taught, including writing, recording and recalling facts.

Weeks 5 - 7

1.4 Follow-up

This was conducted over a total of 3 weeks after the pupils had returned to school. The aim of follow-up was to enhance generalisation and maintenance of behavioural change. Each pupil was visited in school by an Endeavour Centre staff member over a 2 week period. In addition, the teachers of each child completed the teacher rated behaviour card for a total of 2 weeks.

The major objectives of the personal and social development programme were to promote: (1) group skills and peer relations; (2) self-esteem; (3) concentration; (4) rational problem solving skills, including alternative and consequential thinking; (5) self-control and discipline skills; (6) responsibility for self and one's behaviour.

Techniques used included:

(1) role plays;

(2) modelling;

(3) video;

(4) discussion groups;

(5) relaxation;

(6) token economy;

(7) social reinforcement (praise, body contact);

(8) behavioural contingencies;

(9) response cost;

(10) time out;

(11) individual support sessions;

(12) parent programme.

RESULTS AND DISCUSSION

Sample

141 pupils were nominated. Using the randomisation procedure, 70 experimental and 71 control pupils were selected. Of the 70 experimental pupils, 65 (93%) consented to participate in the study and Endeavour Centre programme. Non-consenters included four high school boys and one primary school boy. 61 (94%) completed the programme; they comprise the experimental group. Pupils who failed to complete the programme included one primary school boy, one high school boy and two high school girls.

Of the 71 control group pupils, consent was obtained for participation by 62 (87%) pupils and their parents. Non-consentors included five high school boys. Contact was not made with parents of four pupils. These pupils were subsequently not included in collection of pupil or parent measures. Consent was also sought from the schools for participation of teachers. Consent was obtained for participation of the teachers of 66 pupils (93%). The demographic

characteristics of the sample are demonstrated in Table 7. Analysis was conducted using the BMDP (4F) statistical package (11).

TABLE 7

Demographic Characteristics of the Experimental and Control Groups

		Experimental (n=53)	Control (n=26)	Total (n=79)
Age (%)	10	22.6	19.2	21.5
	11	20.8	26.9	22.8
	12	22.6	23.1	22.8
	13	20.8	19.2	20.3
	14	13.2	11.5	12.7
Sex (%)		(n=61)	(n=66)	(n=127)
Male		80.3	71.2	75.6
Female		19.7	28.8	24.4
Ethnicity (%)		(n=48)	(n=26)	(n=74)
Anglo Saxon Australian		100.0	92.3	97.3
Aboriginal Australian		0.0	3.8	1.4
Other		0.0	3.8	1.4
Single Parent (%)		(n=52)	(n=26)	(n=78)
Home		36.5	19.2	30.8
* Father's Occupation (%)		(n=44)	(n=21)	(n=65)
A (highest status)		0.0	0.0	0.0
B		6.8	4.8	6.2
C		29.5	38.1	32.3
D (lowest status)		27.3	33.3	29.2

Father's Occupation continued (%)	(n=44)	(n=21)	(n=65)
Home Duties	0.0	0.0	0.0
Unemployed	6.8	0.0	4.6
Pensioner	18.2	9.5	15.4
Retired	0.0	0.0	0.0
Student	0.0	0.0	12.3
Non Clasifiable	11.4	14.3	0.0

Mother's Occupation (%)	(n=50)	(n=26)	(n=76)
A	0.0	0.0	0.0
B	6.0	3.8	5.3
C	4.0	11.5	6.6
D	10.0	15.5	11.8
Home Duties	62.0	69.2	64.5
Unemployed	0.0	0.0	0.0
Pensioner	12.0	0.0	7.9
Retired	0.0	0.0	0.0
Student	0.0	0.0	3.9
Not Classifiable	6.0	0.0	0.0

* re Congalton, 1969 (44)

Analyses

Preliminary repeated measures analyses of variance indicated significant pre-intervention differences in the outcome measures between the experimental and control groups. To overcome this problem, subsequent analyses were t-tests performed on Time 1 - Time 2 and Time 1 - Time 3 difference/scores. Analyses compared pre-intervention (Time 1) with one-month follow-up (Time 2) plus

pre-intervention (Time 1) with 6 month follow-up (Time 3). Analyses of outcome measures were conducted using SAS statistical software (43).

(i) Parent Measures. Both parents/guardians were asked to complete the Achenbach Child Behaviour Checklist. Because of the high proportion of single parents (30%), only one measure was analysed for each pupil. These included responses from a mother or stepmother. If this was unavailable, responses from male parents/guardians were used. The change in the social competence score was not significantly different for the two groups for the two time periods (pre-intervention versus 1 month follow-up: $t = 0.80$, $df = 20$, $p>0.05$; pre-intervention versus 6 month follow-up; $t = 0.59$, $df = 36$, $p>0.05$). Similarly, non significant results were noted for the total behaviour score (pre-intervention versus 1 month follow-up: $t = 1.34$, $df = 28$, $p>0.19$; pre-intervention versus 6 month follow-up: $t = 0.40$, $df = 29$, $p>0.05$). Thus, experimental group parents did not report an improvement in childrens' social competence or a decrease in negative behaviour.

(ii) Pupil Measures - Academic Scores. Results indicated that the change in the comprehension percentile rank was not significantly different for the experimental and control groups for the two time periods. (Pre-intervention versus 1 month follow-up: $t = 0.19$, $df = 92$, $p>0.05$). (Pre-intervention versus 6 month follow-up: $t = 1.06$, $df = 80.4$, $p>0.05$). Thus, comprehension scores for the experimental group do not show significant improvement compared with the control group.

The change in the maths percentile rank was not significantly different for the two groups in the pre-intervention versus one month follow-up phase ($t = 0.12$, $df = 96.2$, $p>0.05$). This was also the case for the pre-intervention versus 6 month phase. However, the trend was in the expected direction indicating an improvement in the experimental group ($t = -1.17$, $df = 76.4$, $p>0.05$).

Lack of significant change might be explained by examining pre-intervention academic scores. Results indicate that 46% of the experimental and 60% of the control group had pre-intervention comprehension scores in the lowest 3 stanines. Similarly for maths, 39% of the experimental group and 59% of the control group performed in the lowest 3 stanines. This indicates that a substantial proportion of the pupils did not meet the criteria of being of average and above ability. Rather, it appears that the majority were in need of remedial intervention which is not the aim of the programme.

A number of studies indicate the importance in cognitive-behavioural therapy of targetting specific skill training to the developmental level of the children. While developmental level is not necessarily the same as academic level, the two are often inter-related. Furman (45) reviewed evidence that suggests that children's responsiveness to certain reinforcement and punishment techniques, as well as their responsiveness to certain components of modelling procedures, change developmentally. In addition, Schleser, Meyers and Cohen (46) indicated that developmental status affected self-instructional interventions.

A review of cognitive-behavioural interventions in academic skills by Wong (47) also stresses the importance of pre-skills. Pre-skills are pre-requisites for the development of academic skills. It is noted that, if the necessary pre-skills are inadequate, there may be difficulty in achieving improvement in academic skills.

The results also indicate a non-significant trend for the experimental pupils to improve in relation to the control group from pre-intervention to 6 month follow-up in maths skills. Thus, while a change was not noted immediately, some improvement was indicated at a later stage. It may be the case that some academic skills need time to consolidate. In this case, they will show up progressively in successive delayed post-tests, but not in immediate post-tests (47). This may be the case with maths skills. A 12-month

follow-up would test this hypothesis.

Self-reported Behaviour

The change in the total problem behaviour score on the self-report form was not significantly different for the two groups from pre-intervention to one month follow-up ($t = -0.07$, $df = 73.0$, $p>0.05$). For the pre-intervention versus 6 month follow-up phase, however, a significant decrease in problem behaviour was noted for the control group compared with the experimental group ($t = -2.30$, $df = 56.8$, $p<0.05$). This indicates an improvement in self-reported behaviour in the control group.

Other reports on cognitive-behavioural interventions have also noted an improvement in the control group (48). There are two possible explanations for these results. First, a "spillover effect" may have occurred (49). This refers to changes in the behaviour of control children after an intervention has produced a change in the behaviour of experimental children. The most frequent explanation of the effect is in terms of observational learning (50). An early study by Cooke and Apolloni (51) found control children increased social interaction rates subsequent to the occurrence of training procedures for experimental children. The occurrence of a "spillover effect" seems unlikely in this study, however, given the non-significant change in the experimental group. A more reasonable explanation is that the testing sessions serve as an intervention in themselves for control children. This is possible considering the sessions lasted about a half-day and offered the opportunity for more personal interactions with the tester.

(iii) Teacher Measures - Achenbach Teachers' Report Form. The Achenbach Teachers' Report Form was completed by English teachers for high school pupils and class teachers for primary school pupils. Three scores were analysed, including an adaptive functioning score, school performance score and problem behaviour score.

First, the change in the adaptive functioning score was not significantly different for the two groups for the two time periods. (Pre-intervention versus 1 month follow-up: $t = -0.25$, $df = 51.1$, $p>0.05$). (Pre-intervention versus 6 month follow-up: $t = -0.88$, $df = 41.3$, $p>0.05$). Thus teacher-reported positive behaviour did not increase for the experimental group at a greater rate than for the control group.

Similar results were indicated for the school performance scores. (Pre-intervention versus 1 month follow-up: $t = 0.22$, $df = 43.8$, $p>0.05$). (Pre-intervention versus 6 month followo-up: $t = -0.49$, $df = 39.9$, $p>0.05$). Thus teachers did not report experimental pupils as improving in school performance compared with control pupils.

A significant increase was noted in the total problem behaviour score of the experimental group compared with the control group for the pre-intervention versus 1 month follow-up phase ($t = 2.33$, $df = 42.2$, $p<0.05$). This indicates that teachers reported a short-term increase in problem behaviour in the experimental group. This does not extend to the 6 month follow-up period however, where no significant difference was noted for the two groups ($t = 0.47$, $df = 27.1$, $p>0.05$).

Thus, the experimental group did not improve at a greater rate than the control group in teacher reported measures of academic performance and overall behavioural functioning. In addition, a short-term deterioration was indicated in reported behaviour in the experimental group. Similar results have been noted by Abikoff and Gittelman (52), who examined cognitive behavioural therapy plus medication treatment for hyperactive children. They report that behavioural deterioration following treatment was serious enough during a 1-month follow-up period to require medication.

There are two possible explanations for the results of the present study. First, methods employed to enhance generalisation to the school setting may be

insufficient to prepare the pupil's transition from the Endeavour Centre to school. These methods include the pupil's return to school on Friday to practice skills learnt through the week, plus the use of conduct books and visits by Endeavour Centre staff for three weeks after completion of the programme. Even though these methods seem quite intensive, it could be that the pupils are insufficiently prepared for return to the school system. It would be expected that this transition would be difficult, given the more personalised nature of interactions at the Endeavour Centre compared to those at school. Thus, they may return to school with unrealistic expectations or be unprepared for the stark difference between the Endeavour Centre system and school environment. In response, they may behave more negatively.

An alternative explanation is that teachers are unprepared for the pupil's return. It may be that they hold unrealistic expectations about the pupil's behaviour and expect a total transformation.

Endeavour Behaviour Rating Scale

Three scores were obtained from the Endeavour Behaviour Rating Scale including a total behaviour score, opinion score and description score. The total behaviour score was obtained by using the 26 reliable items on the 5-point rating scale. Totals with more than 2 items missing were excluded from the analysis. A reduction in the score would indicate that teachers reported a decrease in negative behaviour. The opinion score was obtained by asking teachers to rate the pupil in comparison with peers for degree of behavioural disturbance. A reduction in the score would indicate that teachers viewed the pupil as less behaviourally disturbed. The description score was obtained by adding teachers' responses to a set of 10 general descriptors (e.g. aggressive, withdrawn) of behaviour. As with the other two scores, a decrease indicates a reported reduction in negative behaviour.

First, the response from English and class teachers will be reported. The change in the three scores was not significantly different for the two groups for the two time periods. This means that teachers did not report experimental pupils as decreasing in negative behaviour when compared with control pupils.

Results are shown in Table 8.

Table 8: Statistics of English and Class Teachers Responses on the Endeavour Behaviour Rating Scale

	Time	
Score (*p>0.05)	Pre vs 1 month follow-up	Pre vs 6 month follow-up
Total behaviour	t = 1.58, df = 64*	t = 1.05, df = 54*
Opinion	t = 0.79, df = 49*	t = 0.75, df = 47*
Description	t = 0.45, df = 62*	t = 1.92, df = 47*

Similar results were indicated for maths teachers' responses. The change in the three scores was not significantly different for the two time periods. Thus, maths teachers also failed to report a decrease in negative behaviour in experimental pupils. Statistics are shown in Table 9.

Table 9: Statistics of Maths Teachers' Responses on the Endeavour Behaviour Rating Scale

	Time	
Score (*p>0.05)	Pre vs 1 month follow-up	Pre vs 6 month follow-up
Total behaviour	t = 1.92, df = 40*	t = 1.54, df = 10*
Opinion	t = 0.78, df = 38*	t = 0.25, df = 17*
Description	t = 0.60, df = 37*	t = 0.78, df = 12*

Results of the pre-intervention versus 1-month follow-up period indicate a decrease in the total behaviour score in the experimental group. This trend approaches statistical significance ($p=0.06$). Thus in the short term, experimental pupils tend to exhibit a greater reduction in negative behaviour compared with control pupils.

Conclusions and Implications

The methodology employed in the present study is an improvement on that used in previous research, which has been characterized by lack of control groups and follow-up testing and the use of non-clinical samples (36,37). In addition the examined intervention aimed at improving children's overall sense of competence, differing from previous interventions which targetted explicit and isolated skills (53).

The current research however, appears to support previous studies which fail to indicate global improvements in maintenance and generalisation as a result of cognitive behaviour therapy (36).

To summarise the results: First, parents of experimental children did not reported a non-significant improvement in social competence compared with control group parents. Similarly, a non-significant decrease in problem behaviour was noted.

Second, while there was no statistically significant improvement in experimental group academic scores there was a non-significant trend for the experimental group to improve in comparison to the control group at 6 month follow-up in maths skills. No significant improvement relative to the control group was noted for self-reported behaviour. Thus for pupil measures, while a minor improvement is noted, this does not generalise to both academic and behavioural domains.

Results from the Teacher Report Form fail to indicate an improvement relative to the control group in teachers perceptions of school performance and general functioning. Alternatively, a short term deterioration in behaviour relative to that of the control group is noted.

Similarly, for the Endeavour Rating Scale, Maths and English teachers do not improve their opinion of the pupils' behaviour relative to the control group nor describe pupils in less negative terms. In addition, there was no significant difference between groups in negative behaviour as indicated in the total behaviour score. However, there was a non-significant trend for the

behaviour scores given by Maths teachers to the experimental group to be better than those given to the control groups.

Thus, experimental group pupils' general behaviour (as measured by the Teachers' Report) appears to be worse in the short term than that of the control group. Alternatively a non-significant improvement is noted in the short term in more specific conduct (as measured by the Endeavour Behaviour Rating Scale total behaviour score). Again, the behavioural change does not generalise to broad behaviours.

These results suggest a number of modifications which might be incorporated into future cognitive behaviour interventions for conduct disordered pupils. These are listed below.

1. Firstly, it is important that selection criteria for participants in intervention programmes be established and then followed. It is suggested that pupils be tested before entry to ensure that those who fail to meet criteria can be excluded and referred to more suitable agencies.

2. It is suggested that teachers and parents be involved in the program as much as possible to facilitate generalisation and maintenance. The importance of incorporating environmental support is suggested by the research literature (25). While a parent training programme was incorporated into the Endeavour Centre programme, no comparable teacher programme was conducted.

3. It is recommended that the Endeavour Programme be extended to a longer intervention. While there are advantages of short-term intervention (54), there is some suggestion that lengthier treatments have more successful outcome (54).

4. It is recommended that interventions incoporate intensive programming for generalisation. While the Endeavour Centre aimed to achieve this, the results suggest that in the short term, pupils were not prepared for return to school.

References

1. Desk Reference to the Diagnostic Criteria from DMS III, Washington DC, American Psychiatric Association, 1982.

2. Achenbach, T.M. Assessment and taxonomy of child and adolescent psychopathology. Volume 3, California: Sage Publications, 1985.

3. Verhulst, F.C., Akkerhuis, G.W. Mental health in Dutch children (III): behavioural-emotional problems reported by teachers of children aged 4-12. Acta Psychiatrica Scandinavica, 1986, 74, Supplementum No. 330.

4. Kirschenbaum, D.S., Ordman, A.M. Preventive interventions for children: cognitive behavioural perspectives. In Meyers, A.W., Craighead, W.E. (eds), Cognitive Behaviour Therapy with Children, New York, Plenum Press, 1984.

5. Offord, D.R. Prevention of behavioural and emotional disorders in children. Journal of Child Psychology and Psychiatry, 1987, 28, 1, 9-19.

6. Rutter, M. & Giller, H. Juvenile Delinquency Trends and Perspectives, New York, Guilford Press, 1984.

7. Sanson-Fisher, R.W., Martin, C.J. Standardized interviews in psychiatry: Issues of reliability. British Journal of Psychiatry, 1981, 139, 138-143.

8. Maclure, M., Willett, W.C. Misinterpretation and misuse of the kappa statistic. American Journal of Epidemiology, 1987, 126 (2), 161-169.

9. Lee, P.S., Suen, H.K. The estimation of kappa from percentage agreement interobserver reliability. Behavioural Assessment, 1984, 6 (4), 375-378.

10. Cohen, J. Weighted Kappa: Nominal scale agreemnt with provision for scaled disagreement of partial credit. Psychological Bulletin, 1968, 70, 213-20.

11. Dixon, W.J. (ed). BMDP Statistical Software 1981. Berkeley: University of California Press, 1981.

12. Johnson, D.S. The organizing role of motivation in health beliefs and intention. Journal of Health and Social Behaviour, 1972, 13, 285-293.

13. Gorsuch, R.L. Factor analysis. Second edition. New Jersey: Lawrence Erlbaurn Assocates, 1983.

14. Hair, J.F., Anderson, R.E., Tatham, R.L. & Grablowsky, B.J. Multivariate data analysis with readings. Oklahoma: Petroleum Publishing Co., 1979.

15. Pascoe, G.C. & Attkisson, C.G. The evaluation ranking scale: a new methodology for assessing satisfaction. Evaluation and Program Planning, 1983, 6, 335-347.

16. Goyette, C.H., Conners, C.K. & Ulrich, R.F. Normative data on revised Conners Parent and Teacher Rating Scales. Journal of Abnormal Child Psychology, 1978, 6, 221-236.

17. Quay, H. A dimensional appraoch to behaviour disorder. The Revised Behaviour Problem Checklist. School Psychology Review, 1983, 12 (3), 244-249.

18. Edelbrock, C., Greenbaum, R. & Conover, N.C. Reliability and concurrent relations between the teacher version of the Child Behaviour Profile and the Conners Revised Teacher Rating Scale. Journal of Abnormal Child Psychology, 1985, 13 (2), 295-304.

19. Edelbrock, C. & Achenbach, T.M. The teacher version of the Child Behaviour Profile: I. Boys aged 6-11. Journal of Consulting and Clinical Psychology, 1984, 52 (2), 207-217.

20. Gresham, F.M. Utility of cognitive - behavioural procedures for social skills training with children: a critical review. Journal of Abnormal Child Psychology, 1985, 13 (3), 411-423.

21. Gresham, F.M. & Lemanek, K.L. Social skills: A review of cognitive-behavioural training procedures with children. Journal of Applied Development Psychology, 1983, 4, 239-261.

22. Achenbach, T.M. & Edelbrock, C. (1986). Manual for the Teacher's Report Form and Teacher Version of the Child Behaviour Profile. Burlington, VT: University of Vermont Department of Psychiatry.

23. Stott, D.H. Bristol Social Adjustment Guides in D.J. Keyser, R.G. Sweetland (eds). Test Critiques. Missouri: Test Corp. of America, 1984.

24. Achenbach, T.M. & Edelbrock, C. Child Behaviour Checklist. In D.J. Keyser and R.G. Sweetland (eds). Test Critiques, Vol 1. Missouri: Test Corp. of America, 1984.

25. Craighead, W.E., Meyers, A.W. & Craighead, L.W. A conceptual model for cognitive-behaviour therapy with children. Journal of Abnormal Child Psychology, 1985, 13 (3), 331-342.

26. Nunnally, J.C. Introduction to psychological measurement. New York: McGraw-Hill Inc., 1970.

27. Nunnally, J.C. Psychometric theory. New York: McGraw, Hill, 1978.

28. Helmstader, G.C. Principles of psychological measurement. New York: Appleton-Century-Crofts, 1964.

29. Cockburn, J., Fahey, P. & Sanson-Fisher, R.W. (1987). Construction and validation of a questionnaire to measure the health beliefs of general practice patients. Family Practice, 4 (2), 108-116.

30. Connell, H.M. Essentials of Child Psychiatry, 2nd edition. Oxford: Blackwell Scientific Publications, 1985.

31. Rutter, M. Psychopathology and development: 1. Childhood antecedents of adult psychiatric disorder. Australian and New Zealand Journal of Psychiatry, 1984, 18, 225-234.

32. Spivack, G., Marcus, J. & Swift, M. Early classroom behaviours and later misconduct. Developmental Psychology, 1986, 22, 1, 124-131.

33. Galloway, D. Disruptive Pupils. Wellington, New Zealand Council for Educational Research, set 1981, No.1, 1981.

34. Wells, P. The need for mental health services for adolescents in the Hunter Region. Australian and New Zealand Journal of Psychiatry, 1981, 15, 355-359.

35. Casey, R.J. & Berman, J.S. The outcome of psychotherapy with children. Psychological Bulletin, 1985, 98, 2, 388-400.

36. Hobbs, S.A., Moguin, L.E., Tryroler, M. & Lahey, B.B. Cognitive behaviour therapy with children: Has clinical utility been demonstrated? Psychological Bulletin, 1980, 87, 1, 147-165.

37. Urbain, E.S. & Kendall, P.C. Review of social-cognitive problem-solving interventions with children. Psychological Bulletin, 1980, 88, 1, 109-143.

38. Achenbach, T.M. & Edelbrock, C. (1983) Manual for the Child Behaviour Checklist and Revised Child Behaviour Profile. Burlington, VT: University of Vermont Department of Psychiatry.

39. Australian Council for Educational Research. Progressive achievement tests in reading comprehension and vocabulary. Teachers Handbook. Second edition. Hawthorn: Australian Council for Educational Research Limited, 1986.

40. Australian Council for Educational Research. PAT Maths Progressive Achievement Tests in Mathematics. Teachers Handbook. Victoria: Australian Council for Educational Research Ltd, 1984.

41. Achenbach, T.M., Edelbrock, C. Youth self-report - for ages 11-18. Burlington: University of Vermont, 1981.

42. McConaughy, S.H. Using the Child Behaviour Checklist and related instruments in school-based assessment of children. School Psychology Review, 1985, 14 (4), 479-494.

43. Hellwig, J.T., Council, K.A. SAS Users-guide, 1979. North Carolina: SAS Institute Incorporated, 1979.

44. Congalton, A.A. Status and prestige in Australia. Melbourne: FW Cheshire Publishing Pty. Ltd., 1969.

45. Furman, W. Promoting social development. Developmental implications for treatment. In B.B. Lahey and A.E. Kazdin (Eds)., Advances in clinical child psychology, Volume 3. New York: Plenum, 1980.

46. Schleser, R., Meyers, A.W. & Cohen, R. Generalization of self-instructions. Effects of general versus specific content, active rehearsal and cognitive level. Child Development, 1981, 52, 335-340.

47. Wong, B.Y.L. Issues in cognitive-behavioural interventions in academic skill areas. Journal of Abnormal Child Psychology, 1985, 13, 3, 425-442.

48. O'Leary, K.D. & Carr, E.G. Childhood Disorders. In Wilson, G.T., Franks, C.M. (eds). Contemporary Behaviour Therapy. New York: Guilford Press, 1982.

49. Strain, P.S., Shores, R.E. & Kerr, M.M. Experimental analysis of spillover effects on social interaction of behaviourally handicapped preschooler children. Journal of Applied Behaviour Analysis, 1976, 9, 34-40.

50. Bandura, A. Social learning theory. Englewood Cliffs, New Jersey: Prentice Hall, 1977.

51. Cooke, T.P. & Apolloni, T. Developing positive social-emotional behaviours: a study of training and generalization effects. Journal of Applied Behaviour Analysis, 1976, 9, 65-78.

52. Abikoff, H., Gittelman, R. Cognitive training effects in hyperactive children maintained on stimulants. In. Whalen, C.K., Henker, B., Hinshaw, S.P. Cognitive-behavioural therapies for hyperative children: premises, problems and prospects. Journal of Abnormal Child Psychology, 1985, 13 (3), 391-410.

53. Mahoney, M.J. & Nezworski, M.T. Cognitive-behavioural approach to children's problems. Journal of Abnormal Child Psychology, 1985, 13 (3), 467-476.

54. Lochman, J.E., Lampron, L.B., Burch, P.R. & Curry, J.F. Cleint characteristics associated with behaviour change for treated and untreated aggressive boys. Journal of Abnormal Child Psychology, 1985, 13 (4), 527-538.