THE RUTTER SCALE FOR COMPLETION BY TEACHERS: FACTOR STRUCTURE AND RELATIONSHIPS WITH COGNITIVE ABILITIES AND FAMILY ADVERSITY FOR A SAMPLE OF NEW ZEALAND CHILDREN

ROB MCGEE,* SHEILA WILLIAMS,* JENNIFER BRADSHAW,† JAMES L. CHAPEL,‡ ARTHUR ROBINS‡ and PHIL A. SILVA*

*University of Otago, Dunedin, †Department of Justice, Dunedin and ‡University of Missouri—Columbia

Abstract—A large sample of 7-year-old children ($n = 940$) was rated by teachers using the Rutter Child Scale B, a 26-item questionnaire covering a variety of behavioural problems. A factor analysis of the data revealed three main factors of interest, identified as aggressiveness, hyperactivity and anxiety–fearfulness. Measures based upon these factors had a reasonably high level of reliability and were moderately stable over a 2-year interval. An analysis of the relationship between these three behavioural measures and some cognitive measures indicated that only hyperactivity was negatively associated with cognitive ability. However, both hyperactivity and aggressiveness were related to adversity in the child's family background. The findings suggest the usefulness of distinguishing between aggressive and hyperactive dimensions of behaviour.

Keywords: Teacher Rutter Scale, factor structure, cognitive ability, family adversity

INTRODUCTION

THE Children's Behaviour Questionnaires for completion by parents and teachers (Child Scales A and B: Rutter, 1967; Rutter, Tizard & Whitmore, 1970) were designed to provide valid and reliable screening measures of a child's behaviour at home and at school. These questionnaires have been used by Rutter and his associates in large-scale epidemiological surveys in the Isle of Wight (Rutter *et al.*, 1970) and Inner London (Rutter, Cox, Tupling, Berger & Yule, 1975). Recently, Venables, Fletcher, Dalais, Mitchell, Schulsinger & Mednick (1983) have reported some results on the use of the teacher Rutter questionnaire with a large sample of children on the island of Mauritius.

Both of the Rutter scales have been administered to the parents and teachers of children enrolled in the Dunedin Multidisciplinary Child Development Study. This

*The Dunedin Multidiscipilinary Health and Development Research Unit, Department of Paediatrics and Child Health, Medical School, University of Otago, Dunedin, New Zealand.
†Department of Justice, Dunedin, New Zealand.
‡Department of Psychiatry, University of Missouri—Columbia, MO, U.S.A.
Requests for reprints to: Dr. Rob McGee, The Dunedin Multidisciplinary Health and Development Research Unit, Department of Paediatrics and Child Health, Medical School, University of Otago, Box 913, Dunedin, New Zealand.

is a longitudinal study of the health, development and behaviour of a large sample of children from Dunedin, New Zealand [see McGee & Silva (1982) for a summary of the characteristics and history of the sample]. Some of the initial results from the use of the scales with this sample are described by McGee, Silva & Williams (1983, 1984a). The aims of the present paper are (1) to provide information concerning the psychometric properties of these scales; (2) to examine the strength of the relationships between the behaviour ratings and the children's cognitive development; and (3) to compare the results obtained from the New Zealand sample with those from samples of children in other countries.

The Rutter scales consist of a series of behavioural items to which the respondent replies: "does not apply"—scored 0; "applies somewhat"—scored 1; and "certainly applies"—scored 2. The parent and teacher forms sample essentially the same pool of items, although parents are asked to report on a number of specific behaviours in a little more detail. In the original research describing the use of these scales, a total score was obtained by summing individual item scores. Cut-off scores were then used to identify children with high levels of reported problem behaviours. For example, in the case of the teacher scale a total score of 9 or more indicated a child likely to show some behavioural or emotional disorder. Once such a child was identified, 'neurotic' and 'antisocial' subscales were obtained so that the nature of the child's problems could be described. The power of the scale to discriminate between these two types of problems was tested by comparing questionnaire results with diagnoses made from the case notes of clinic children (see Rutter, 1967; Rutter et al., 1970).

In subsequent years there have been relatively few publications concerning the psychometric properties of these scales. There appear to be four papers describing the results of factor analyses of the questionnaires. Venables et al. (1983) reported a factor analysis of the teacher scale used with a sample of over 1000 7- to 8-year-olds. They described a two-factor solution, with one factor characterised as 'aggressive–hyperactive' and the other 'anxious–fearful'. Schachar, Rutter & Smith (1981), in a re-analysis of the original Isle of Wight data, found that a 'hyperkinetic' factor emerged separately from aggressiveness in a series of principle component analyses. The hyperkinetic factor was made up of 'restless', 'squirminess' and 'cannot settle/poor concentration'. A further two analyses have used a preschool version of the Rutter. Behar & Stringfield (1974), using a modification of the teacher scale with 496 children aged 3–6 years, obtained a three-factor solution with the factors identified as 'hostile–aggressive', 'anxious–fearful', and 'hyperactive–distractible'. Fowler & Park (1979), studying a 28-item version of the preschool questionnaire, reported a two-factor solution with hyperactive items forming part of the aggressiveness factor, although they did find that the hyperactivity items clustered together in a three-factor solution. However, Fowler and Park argued on other grounds that the two-factor solution is more appropriate.

The nature of the factor structure of these scales is of some interest in view of the debate regarding the status of 'hyperactivity' as a 'disorder' (Quay, 1979; Shaffer & Greenhill, 1979). For example, do hyperactive behaviours need to be distinguished from behaviours usually considered as indicative of conduct disorder? The first part of this paper reports the factor structure of the Teacher Rutter Scale for a large sample

of 7-year-old children in Dunedin. Having identified the underlying dimensions of the scale, it is possible to test the meaningfulness of these dimensions by seeing whether they are differentially related to other classes of variables (Achenbach, 1981). In the second part of the paper the relationships between the behavioural dimensions suggested by the factor analysis and general measures of cognitive ability and family adversity are investigated.

METHOD

Sample

The sample consisted of 951 children (492 boys and 459 girls) aged 7 years who were enrolled in Dunedin Multidisciplinary Child Development Study. This sample has been described by McGee & Silva (1982). In summary, the children were part of a cohort born at Queen Mary Hospital, Dunedin between 1 April 1972 and 31 March 1973. When the sample was first traced at age 3 years, 1139 children lived in the Dunedin metropolitan area or the province of Otago and were thus eligible for inclusion in the study. Of these 1139 children, 1037 or 91% were assessed at 3 years of age; the remaining 9% were either traced too late for inclusion at 3 years or the parents were unable to co-operate. The sample has been assessed thereafter at 2-yearly intervals. The sample itself, when compared with New Zealand as a whole on an index of socio-economic status (Elley & Irving, 1972), tended to be socio-economically advantaged, being over-represented in the two highest levels and under-represented in the two lowest. Furthermore, the sample was under-representative of Maori and Polynesian races: about 2% were more than half Maori and Polynesian compared with about 10% for the whole of New Zealand (*New Zealand Yearbook*, 1976).

Measures

The Rutter Scale B for completion by teachers (Rutter, 1967) consists of 26 statements. These teacher ratings were initially collected during the children's first year at school and again at age 7 years. At the latter age a total of 940 questionnaires (491 boys and 449 girls) were available for analysis.

Three measures were used as indices of the child's current cognitive ability. These were the verbal intelligence (VIQ) and performance intelligence (PIQ) scales of the WISC-R measure of intelligence (Wechsler, 1974) and reading ability measured by the Burt Word Reading Test—1974 Revision (Scottish Council for Research in Education, 1976). These measures were part of a comprehensive health, developmental and behavioural assessment completed when the children enrolled in the Dunedin study attended the research unit. Children living in other parts of New Zealand were assessed by the Psychological Service in the Department of Education. In the majority of cases, each child was assessed within one month of their seventh birthday. The psychometrists carrying out the cognitive assessments had no knowledge of the behavioural questionnaire results.

In order to study the relationship between behaviour and family background, it was decided to use a 'family adversity index' of the kind reported by Rutter (1978). It should be noted, however, that while using Rutter's approach the present index differs from that reported by Rutter in several key features. For example, the Dunedin index does not include whether the child has ever been 'in care', or whether the child's father has committed an offence against the law. The index developed was based upon (1) SES levels V and VI, representing semi-skilled and unskilled father's occupation; (2) large family size of four or more children; (3) parental separations up to age 7 years; (4) low maternal mental ability as measured by the SRA verbal test (Thurstone & Thurstone, 1973)—a score of 21 or less identified 11.8% of the mothers; (5) maternal reports of poor 'mental health', suggested by a score of 6 or more on the Malaise Inventory of Rutter *et al.* (1970); and (6) reportedly 'poor' family relationships, indicated by a score of 14 or less on the measures of family cohesion, expressiveness and conflict of the Family Environment Scale (Moos, 1974). Items were scored 1 if present and 0 if absent. In the case of missing values, a conservative approach of scoring the item 0 was adopted. This index has been reported by McGee, Silva & Williams (1984b). A score of 3 or more selected 8% of 940 children who might be considered to be from adverse family circumstances.

RESULTS

Item distributions

Table 1 shows the distribution of responses for each of the 26 items at age 7 years for boys and girls separately. There was considerable variation in the distribution of scores among the items. Truancy was the least frequently occurring behaviour, as might be expected at this age. Poor concentration was the most frequently noted behaviour among boys, while worrying was the most frequently noted behaviour in girls. Table 1 also shows those items having significantly different distributions according to sex. Boys were more frequently described as restless, squirmy, having poor concentration, disobedient and, to a lesser degree, getting into fights and being more irritable than girls. Girls, on the other hand, were noted as more often complaining of aches and pains than boys.

Factor analysis

In order to examine the structure of the teacher questionnaire, initial factor analyses were performed on the data from the boys and girls separately, using

TABLE 1. ITEM DISTRIBUTIONS FOR BOYS AND GIRLS ON THE RUTTER CHILD SCALE B

Item	Boys (n = 491)			Girls (n = 449)			χ^2 (2 d.f.)
	Does not apply	Applies somewhat	Certainly applies	Does not apply	Applies somewhat	Certainly applies	
Restless	67.0	26.9	6.1	82.6	14.0	3.4	30.12***
Truants	98.8	1.0	0.2	99.6	0.4	0.0	1.96
Squirmy	66.6	26.5	6.9	82.8	14.3	2.9	32.92***
Destructive	92.1	6.9	1.0	95.1	4.2	0.7	3.59
Fights	77.6	18.9	3.5	85.1	13.6	1.3	10.05**
Not liked	85.8	12.6	1.6	84.0	13.6	2.4	1.03
Worries	64.0	32.4	3.6	68.4	27.4	4.2	2.83
Solitary	68.2	26.5	5.3	68.8	28.1	3.1	2.84
Irritable	85.9	11.0	3.1	92.2	6.0	1.8	9.34**
Miserable	88.2	10.4	1.4	84.4	13.4	2.2	2.98
Twitches	94.5	4.1	1.4	95.1	4.0	0.9	<1.00
Sucks thumb	95.1	3.7	1.2	92.6	4.7	2.7	3.30
Bites nails	96.1	2.9	1.0	93.6	5.3	1.1	3.79
School absence	96.3	3.3	0.4	93.8	5.3	0.9	3.42
Disobedient	77.8	18.7	3.5	89.8	8.0	2.2	25.05***
Poor concentration	58.7	28.9	12.4	77.1	18.7	4.2	40.44***
Fearful	72.5	24.4	3.1	71.7	25.2	3.1	<1.00
Fussy	88.4	10.6	1.0	83.5	15.8	0.7	5.87
Lies	88.2	10.2	1.6	91.3	6.7	2.0	3.82
Steals	94.7	3.9	1.4	95.8	3.3	0.9	<1.00
Wets/soils self	96.9	2.1	1.0	97.3	1.8	0.9	<1.00
Aches/pains	95.5	4.5	0.0	87.3	11.1	1.6	22.94***
Tears at school	98.6	1.4	0.0	96.2	2.7	1.1	7.41*
Stutters	94.7	4.3	1.0	97.3	2.2	0.5	4.19
Other speech problem	87.0	8.7	4.3	90.0	7.6	2.4	2.94
Bullies	87.0	10.8	2.2	89.3	9.1	1.6	1.36

Figures shown are percentages.
Distributions show significant sex differences: *$P<0.05$; **$P<0.01$; ***$P<0.001$.

principal factoring with iteration, followed by varimax rotation (Nie, Hull, Jenkins, Steinbrenner & Bent, 1975). A comparison of the factor structures from these analyses indicated little difference between the boys and the girls with factors loading on the same items in each analysis. For the sake of economy of presentation of results, it was decided to re-analyse the results from the sample as a whole. Prior to this re-analysis, three items were excluded from further consideration, namely 'truancy', 'wets/soils' and 'tears on arrival at school'. This was done because of the poor discriminative power of these items; all three had fewer than a 5% frequency of 1 and 2 scores in the case of both boys and girls. These items also showed very low communalities on the separate factor analyses for boys and girls.

The factor analysis of the remaining 23 items yielded six factors with eigenvalues greater than 1. The eigenvalues before rotation were 6.06, 2.42, 1.43, 1.32, 1.12 and 1.06; the eigenvalue of the seventh factor ws 0.95 The first six factors accounted for 58.3% of the variance. Table 2 shows the factor loadings after rotation. Factor 1, loading on destructiveness, fighting, not being liked, irritability, disobedience, lying and bullying, appears to be an 'aggressive–antisocial' factor. Factor II (worries, solitary, miserable, fearful and fussy) is very similar to the Rutter et al. (1970) 'neurotic' grouping of items and the Behar & Stringfield (1974) and Venables et al. (1983) 'anxious–fearful' factor. Factor III, loading on restless, squirmy and poor concentration, represents the same 'hyperactivity factor' as identified by Schachar et al. (1981).[*] The remaining three factors only loaded on pairs of items and as such may be considered indeterminate. Stealing was correlated with lying ($r = 0.56$) but had low correlations with the other behavioural items. Similarly, twitches was the only item having a moderate correlation with stuttering ($r = 0.32$), and thumb-sucking and nail-biting ($r = 0.34$) correlated with little else. Thus the most satisfactory solution appears to be in terms of the first three factors.

For the purposes of further analysis of the relationships among the more significant factors, three measures were formed from the unweighted sums of those items in Factors I to III, shown with an asterisk in Table 2. The item 'lies' was not included in the aggressiveness measure, as it loaded on a second factor (Factor IV). Consequently, each item contributed to only one measure. These three measures were analysed for reliability using coefficient alpha, an index of internal consistency which may be interpreted as the expected correlation between a test and an alternate form of the test with the same number of items (Nunnally, 1967). The values of coefficient alpha were 0.83 for aggressiveness, 0.72 for anxious–fearful and 0.82 for hyperactivity, indicating a reasonably high level of internal consistency for each scale.

Table 3 gives the means and standard deviations for boys and girls for the total score on the teacher questionnaire and for the three measures derived on the basis of the factor analysis. Student's t tests indicated significant ($P < 0.15$) sex differences on the measures of aggressiveness, hyperactivity and total scale score, with boys having higher problem scores than girls. There was no sex difference, however, on the measure of anxiety-fearfulness. Inspection of the distributions for these scores

---

[*]A factor analysis of the 989 teacher questionnaires collected at school entry revealed exactly the same pattern of loadings for the first three factors. "Hyperactivity" emerged as a separate factor from the "aggressiveness" factor.

TABLE 2. SORTED FACTOR LOADINGS FOR ITEMS ON RUTTER CHILD SCALE B AT AGE 7 YEARS ($n = 940$)

Item	Factor†					
	I	II	III	IV	V	VI
Fights	80*	07	20	04	02	06
Bullies	68*	01	15	15	02	08
Irritable	63*	21	16	00	09	06
Disobedient	58*	04	43	22	05	03
Destructive	54*	03	19	36	16	09
Not liked	48*	38	12	18	12	07
Worries	02	74*	00	00	18	11
Fearful	−03	60*	10	11	23	04
Miserable	31	59*	10	15	04	23
Fussy	03	49*	01	−12	02	03
Solitary	11	45*	12	08	09	−05
Squirmy	29	10	77*	10	16	19
Restless	39	−01	72*	06	10	13
Poor concentration	28	15	56*	20	19	06
Lies	50*	03	12	64*	07	04
Steals	27	−02	15	60*	09	11
Twitches	13	14	16	01	54*	16
Stutters	04	05	02	01	51*	07
Sucks thumb	04	06	11	10	08	54*
Bites nails	14	18	07	12	11	50*
Aches/pains	04	33	−05	14	−08	13
School absence	−01	15	04	29	−01	11
Other speech problem	02	13	14	08	21	−04

†Decimal points not shown; factor loadings at or above 0.45 are asterisked. A loading of 0.45 indicates that 20% of the variance on the items is accounted for by the factor.

TABLE 3. MEANS AND STANDARD DEVIATIONS FOR BOYS ($n = 491$) AND GIRLS ($n = 449$) ON THE RUTTER CHILD SCALE B

		Aggressiveness	Measure Hyperactivity	Anxiety–fearfulness	Total score
Boys	Mean	1.1	1.3	1.3	4.5
	S.D.	2.0	1.6	1.7	4.8
Girls	Mean	0.7	0.7	1.4	3.6
	S.D.	1.6	1.3	1.7	4.8
Total	Mean	0.9	1.0	1.3	4.0
	S.D.	1.8	1.5	1.7	4.8

indicated that they were all positively skewed, and as such, this raises the issue of the appropriate type of analysis for data of this kind. Parametric techniques (t and F tests) appear to be robust enough to handle distributions which vary markedly from the normal (McNemar, 1969). In the case of further correlational analyses, rank-order correlations were calculated first and compared with product–moment correla-

tions. Second, the data were transformed using a logarithmic transformation in an effort to normalise scores. As both of these methods yielded essentially the same results as product–moment correlations, the latter are reported in subsequent correlational analyses.

The overall intercorrelations among the three derived behavioural measures were as follows: aggressiveness–hyperactivity, $r = 0.61$; aggressiveness–anxiety/fearfulness, $r = 0.30$; and anxiety/fearfulness–hyperactivity, $r = 0.25$. The partial correlations between these respective item pairs, while controlling for the third behavioural measure, were 0.58, 0.20 and 0.08. All correlations were significant ($P < 0.05$), but clearly the strongest relationship was between aggressiveness and hyperactivity. To examine the stability of these measures over time, aggressiveness and hyperactivity and anxiety–fearfulness scores were calculated in the above manner for the 5-year results. The correlations between the 5- and 7-year scores were 0.46 for aggressiveness, 0.52 for hyperactivity and 0.26 for anxiety–fearfulness, all values $P < 0.05$. The correlation between the total questionnaire scores at both ages was $r = 0.44$, $P < 0.05$.

Behavioural measures and cognitive ability

The overall correlations between the total questionnaire score and the cognitive measures of VIQ, PIQ and reading at age 7 years were −0.23, −0.17 and −0.26, $P < 0.05$ for each correlation respectively. Table 4 gives the partial correlations between each behavioural measure and VIQ, PIQ and reading while controlling for the other two behavioural measures. As initial analyses revealed no significant sex differences in the correlations between the behavioural and cognitive measures, the results are reported for the whole sample ($n = 928$). It is apparent from the table that of the three behavioural measures, only hyperactivity showed significant (but relatively weak) negative correlations with the cognitive scores. Aggressiveness and anxiety–fearfulness were unrelated to cognitive ability.

Behavioural measures and family circumstances

To examine the relationships between the three behavioural measures and adverse family background, the total sample was divided into four groups, scoring 0, 1, 2 and 3 or more on the family adversity index. Table 5 shows the means in each group for the aggressiveness, hyperactivity and anxiety–fearfulness measures, and

TABLE 4. PARTIAL PRODUCT-MOMENT CORRELATIONS BETWEEN THE BEHAVIOURAL MEASURES OF AGGRESSIVENESS, HYPERACTIVITY AND ANXIETY-FEARFULNESS AND READING, VIQ AND PIQ SCORES ($n = 928$)

| Behavioural measures | Cognitive measures | | |
	Reading	VIQ	PIQ
Aggressiveness	0.02	0.00	0.03
Hyperactivity	−0.28*	−0.21*	−0.17*
Anxiety–fearfulness	0.02	0.00	−0.01

*Significant correlation ($P < 0.05$). The correlation between each behavioural and cognitive measure partials out the effects of the two other behavioural measures.

TABLE 5. MEAN SCORES FOR THE MEASURES OF AGGRESSIVENESS, HYPERACTIVITY, ANXIETY-
FEARFULNESS AND TOTAL SCORE ACCORDING TO FAMILY ADVERSITY

| Measure | Family adversity score | | | | $\sqrt{MSe}$† | $F(3,936)$ |
	0 $n = 420$	1 $n = 309$	2 $n = 136$	3 +		$n = 75$
Aggressiveness	0.8	0.9	1.2	1.6	1.82	5.26*
Hyperactivity	0.8	1.0	1.3	1.7	1.49	10.59*
Anxiety–fearfulness	1.3	1.3	1.4	1.6	1.69	1.07 ns
Total	3.4	3.9	4.9	6.5	4.74	10.87*

*$P < 0.05$, otherwise non-significant.

† $\sqrt{MSe}$ represents square root of the mean square error term for ANOVA.

the total score on the teacher questionnaire. One-way analyses of variance indicated that aggressiveness, hyperactivity and total scale scores varied significantly over increasing levels of family adversity. Inspection of the table indicates increases in these problem scores as adversity increased. Anxiety–fearfulness, on the other hand, was not significantly related to family adversity. In order to examine the effects of family adversity in more detail, post-hoc Scheffé comparisons were made contrasting each 'risk' group with the nil adversity goup on the total scale score. The group with one risk factor only did not differ significantly from the group with no risk factors. However, the group with two factors present did differ significantly ($P < 0.05$) from the no risk group, as did the group with three or more factors.

Prevalence estimates of behaviour problems

The final set of results concerns the estimation of prevalence rates of aggressiveness, hyperactivity and anxiety–fearfulness based upon the teacher reports. Glow (1980) has outlined several ways in which prevalence rates of problem behaviours may be estimated from the kind of behaviour reports used in the present study. One approach has been to use a cut-off score of 5, *a priori* (McGee & Silva, 1982). A score of 5 implies that at the very least, three problem behaviours have been reported as present, with two 'certainly applying' to the child. A score of 5 in the case of the three behavioural measures in the present study is similar to that obtained using the mean plus two standard deviation cut-off criteria* employed with such scales as the Conners Teacher Questionnaire (Glow, 1980). Table 6 gives the percentages of boys and girls obtaining a score of 5 or above on these three measures. Aggressive behaviours were the most frequently noted abnormal behaviours in boys, while anxious–fearful behaviours were the most common abnormal behaviour in girls.

A slightly different approach to the prevalence of disorder is a categorical approach (see Glow, 1980) in which individuals are assigned to only one category, e.g. 'pure' hyperactivity. Using this method, children with hyperactive and anxious behaviours would not be classified as hyperactive; the term 'hyperactivity' implies that not only do hyperactive behaviours predominate, but that such behaviours are unaccompanied by other types of problems. Table 7 shows the classification of children scoring 5 or

*Applying the mean plus two standard deviation criteria in the present study results in cut-off scores of 4.5 for aggressiveness, 4.0 for hyperactivity and 4.7 for anxiety–fearfulness.

TABLE 6. PERCENTAGES† OF BOYS AND GIRLS SCORING AT OR ABOVE THE CRITERION SCORE 5 ON THE AGGRESSIVENESS, HYPERACTIVITY AND ANXIETY–FEARFULNESS MEASURES OF THE RUTTER CHILD SCALE B

Measure	Boys	Girls	Boys : girls
Aggressiveness	6.9% ($n = 34$)	5.8% ($n = 26$)	1.2 : 1
Hyperactivity	6.1% ($n = 30$)	2.7% ($n = 12$)	2.3 : 1
Anxiety–fearfulness	5.3% ($n = 26$)	7.1% ($n = 32$)	0.7 : 1

†These percentages are not independent as some children scored 5 on two or three scales.

more on the relevant scales according to this categorical method. In contrast to the preceding approach, each child has been assigned to only one type of disorder. This procedure identified 69 boys (14.0% of all boys) and 51 girls (11.3% of girls), a non-significant difference. However, there was a significant association between type of problem identified and sex of child, χ^2 (5 d.f.) = 12.54, $P < 0.05$. Post-hoc χ^2 comparisons of each problem group with the sample remainder (Everitt, 1977) showed a significant sex difference only in the case of those children reported as hyperactive.

TABLE 7. PREVALENCE OF BEHAVIOUR PROBLEMS USING A CATEGORICAL APPROACH TO DISORDER BASED UPON TEACHER QUESTIONNAIRES

Measure	Boys	Girls	Boys:Girls
Aggressiveness	3.7% ($n = 18$)	2.4% ($n = 11$)	1.5:1
Hyperactivity	3.3% ($n = 16$)	0.4% ($n = 2$)	8.3:1
Anxiety–fearfulness	3.7% ($n = 18$)	5.1% ($n = 23$)	0.7:1
Hyperactive-Aggressive	2.6% ($n = 13$)	2.2% ($n = 10$)	1.2:1
Other—Mixed	0.8% ($n = 4$)	1.1% ($n = 5$)	0.7:1

DISCUSSION

Rutter's (1967) original grouping of items into diagnostic categories was based upon the comparison of individual item scores for neurotic and antisocial children. Thus a child was classified into 'neurotic', 'antisocial' or 'mixed disorder' categories on the basis of scores on the neurotic and antisocial subscales. The findings of the factor analysis in this study support Rutter's groupings of items as 'antisocial' and

'neurotic', and the emergence of separate aggressiveness and anxiety-fearfulness factors agrees with the recent findings of Venables *et al.* (1983). Factor analyses of other behavioural rating scales confirm this distinction between aggressiveness and anxious behaviours (Quay, 1979). However, whereas Venables *et al.* report a two-factor solution for their total sample, the present analysis found that hyperactivity emerged as a separate factor. That this was the case for the factor analysis of boys and girls separately and for the total sample at age 5 years suggests that this is a reliable finding. As such, it confirms the results of Schachar *et al.* (1981) and Behar & Stringfield (1974). These three main factors were made up of only 14 of the total 26 items on the teacher's scale. The remaining items were reported too infrequently to be included in the analysis (for example 'truanting'); occurred in minor factors due to high correlations between individual items ('twitches' and 'stutters'); or did not load on any factors ('other speech problem'). However, it may be the case that as the children grow older, some of these more infrequent problems will occur more often and a factor analysis of data collected during adolescence may result in a different factor pattern.

Behavioural subscales based upon the factor loadings appeared to have a reasonably high degree of reliability. Furthermore, the measures of aggressiveness and hyperactivity showed moderate stability over a 2-year interval from age 5 to age 7 years. Anxiety–fearfulness showed the poorest stability over time. All three measures were significantly intercorrelated, although anxiety–fearfulness was only weakly related to the other two behavioural measures. Aggressiveness and hyperactivity had the strongest association, $r = 0.58$, a level of association comparable to correlations obtained with the Conners Teacher Questionnaire (Lahey, Green & Forehand, 1980). Although there is some debate as to whether hyperactivity and aggressive-antisocial behaviour exist as separate behavioural dimensions, the findings of the present study suggest that the distinction is a meaningful one. First, the two factors emerged separately on a series of factor analyses. Second, measures based upon the factor loadings although correlated reasonably highly, were not perfectly correlated, suggesting that a considerable portion of the variance on the hyperactivity measure was independent of variation on the aggressiveness measure. Third, of the three behavioural dimensions, only hyperactivity was associated with poorer cognitive abilites. Neither aggressiveness nor anxiety-fearfulness *per se* were strongly related to VIQ PIQ and reading once their association with hyperactivity had been controlled. Lastly, children with relatively 'pure' forms of each type of problem behaviour could be identified, suggesting that these three dimensions were independent in a phenomenological sense.

Achenbach (1981), among others, has argued that for a taxonomy of disorders to be useful and meaningful, the behavioural dimensions should be differentially related to other classes of variables, for example, those related to treatment outcome or aetiology. In the present study the three dimensions of aggressiveness, anxiety-fearfulness and hyperactivity were studied in association with both cognitive measures and family adversity. The finding that hyperactivity was most strongly related to cognitive variables suggests its usefulness as an independent behavioural description. Furthermore, while antisocial and anxious–neurotic behaviour have been associated with poor reading ability and school failure (McMichael, 1979), the present results suggest that hyperactive behaviour may be the best predictor of poor reading skills

during early schooling. On the other hand, both hyperactivity and aggressiveness ratings were related to adversity in the child's family environment, while anxious behaviour was not. In terms of overall problem behaviour, the presence of two or more risk factors resulted in a significant increase in level of reported problems compared with children with no adverse family stresses. This finding is in agreement with that reported by Rutter (1978), suggesting an interaction between the risk factors.

Because the Rutter scales have not been widely used outside the United Kingdom, it is difficult to compare the New Zealand findings with other established norms. Furthermore, comparisons with the Isle of Wight findings are made difficult by the age differences in the two samples. In the Dunedin study sex differences were related particularly to hyperactive items (restlessness, squirminess and poor concentration) and antisocial items (fighting, irritability and disobedience), with boys showing these behaviours more often than girls. Girls, on the other hand, were more often described as having aches and pains. For the most part, these sex differences agree with earlier reports (Rutter et al., 1970). When more extreme problem behaviours were examined using cut-off scores on the three subscales, hyperactivity was the only behaviour problem to show a clear sex difference. The finding of a boy:girl ratio for hyperactivity of over 8:1 in the Dunedin sample agrees with results in both Australia (Glow, 1980) and the United States (Sandoval, Lambert & Sassone, 1980).

Finally, it needs to be emphasised that these findings are based upon the teacher's perceptions of children's behaviour. Rutter et al. (1970) have reported a relatively low correlation between teachers' and parents' ratings of the same child. Furthermore, parents' and teachers' perceptions of problem behaviour may be differentially influenced by characteristics of the child and his or her background (McGee et al., 1983). This being so, the present results may not apply to parental ratings of behaviour problems.

REFERENCES

Achenbach, T. M. (1981). The role of taxonomy in developmental psychology. In M. E. Lamb & A. L. Brown (Eds), *Advances in developmental psychology* (Vol. 2, pp. 159–198). Hillside, NJ: Lawrence Erlbaum.

Behar, L. & Stringfield, S. (1974). A behaviour rating scale for the pre-school child. *Developmental Psychology, 10,* 601–610.

Conners, C. K. (1969). A teacher rating scale for use in drug studies with children. *American Journal of Psychiatry, 126,* 884–888.

Elley, W. B. & Irving, J. C. (1972). A socio-economic index for New Zealand based on levels of education and income from the 1966 census. *New Zealand Journal of Educational Studies*, **7**, 155–167.

Everitt, B. S. (1977). *The analysis of contingency tables*. London: Chapman & Hall.

Fowler, P. C. & Park, R. M. (1979). Factor structure of the pre-school behaviour questionnaire in a normal population. *Psychological Reports*, **45**, 599–606.

Glow, R. A. (1980). A validation of Conners TQ and a cross-cultural comparison of prevalence of hyperactivity in children. In G. D. Burrows & J. S. Werry (Eds), *Advances in human psychopharmacology* (Vol. 1, pp. 303–320). Greenwich, CT: JAI Press.

Lahey, B. B., Green, K. D. & Forehand, R. (1980). On the independence of ratings of hyperactivity, conduct problems, and attention deficits in children: a multiple regression analysis. *Journal of Consulting and Clinical Psychology*, **48**, 566–574.

McGee, R. & Silva, P. A. (1982). *A thousand New Zealand children: their health and development from birth to seven*. Special Series Report, No. 8. Auckland: Medical Research Council of New Zealand.

McGee, R., Silva, P. A. & Williams, S. (1983). Parents' and teachers' perceptions of behaviour problems in seven year old children. *The Exceptional Child*, **30**, 151–161.

McGee, R., Silva, P. A. & Williams, S. M. (1984a). Behaviour problems in a population of seven year old children: prevalence, stability, and types of disorder. *Journal of Child Psychology and Psychiatry*, **25**, 251–259.

McGee, R., Silva, P. A. & Williams, S. M. (1984b). Perinatal, neurological, environmental and developmental characteristics of seven-year-old children wtih stable behaviour problems. *Journal of Child Psychology and Psychiatry*, **25**, 573–586.

McMichael, P. (1979). The hen or the egg? Which comes first—antisocial emotional disorders or reading disability? *British Journal of Educational Psychology*, **49**, 226–238.

McNemar, Q. (1969). *Psychological statistics*. New York: Wiley.

Moos, R. (1974). *Combined preliminary manual for the family, work and group environment scales*. Palo Alto, CA: Consulting Psychologists Press.

New Zealand Yearbook (1976). Wellington: New Zealand Government Printer.

Nie, N. H., Hull, C. H., Jenkins, J. G., Steinbrenner, K. & Bent, D. H. (1975). *Statistical package for the social sciences*, 2nd ed. New York: McGraw-Hill.

Nunnally, J. C. (1967). *Psychometric theory*. New York: McGraw-Hill.

Quay, H. C. (1979). Classification. In H. C. Quay & J. S. Werry (Eds), *Psychopathological disorders of childhood* (2nd ed., pp. 1–42). New York: Wiley.

Rutter, M. (1967). A children's behaviour questionnaire for completion by teachers: preliminary findings. *Journal Child Psychology and Psychiatry*, **8**, 1–11.

Rutter, M. (1978). Family, area and school influences in the genesis of conduct disorders. In L. A. Hersov, M. Berger & D. Shaffer (Eds), *Aggression and anti-social behaviour in childhood and adolescence* (pp. 95–133). Oxford: Pergamon Press.

Rutter, M., Cox, A., Tupling, C., Berger, M. & Yule, W. (1975). Attainment and adjustment in two geographical areas—I. The prevalence of psychiatric disorder. *British Journal of Psychiatry*, **126**, 493–509.

Rutter, M., Tizard, J. & Whitmore, K. (1970). *Education, health and behaviour*. London: Longmans.

Sandoval, J., Lambert, N. M. & Sassone, D. (1980). The identification and labelling of hyperactivity in children: an interactive model. In C. K. Whalen & B. Henker (Eds), *Hyperactive children: the social ecology of identification and treatment* (pp. 145–171). New York: Academic Press.

Schachar, R., Rutter, M. & Smith, A. (1981). The characteristics of situationally and pervasively hyperactive children: implications for syndrome definition. *Journal of Child Psychology and Psychiatry*, **22**, 375–392.

Scottish Council for Research in Education (1976). *The Burt word reading test—1974 revision*. London: Hodder & Stoughton.

Shaffer, D. & Greenhill, L. (1979). A critical note on the predictive validity of the hyperkinetic syndrome. *Journal of Child Psychology and Psychiatry*, **20**, 61–72.

Thurstone, T. G. & Thurstone, L. L. (1973). *The SRA verbal form.* Chicago, IL: Science Research Associates.

Venables, P. H., Fletcher, R. P., Dalais, J. C., Mitchell, D. A., Schulsinger, F. & Mednick, S. A. (1983). Factor structure of the Rutter "Children's Behaviour Questionnaire" in a primary school population in a developing country. *Journal of Child Psychology and Psychiatry, 22,* 375–392.

Wechsler, D. (1974). *The Wechsler intelligence scale for children—revised.* New York: Psychological Corporation.

This document is a scanned copy of a printed document. No warranty is given about the accuracy of the copy. Users should refer to the original published version of the material.