

ORIGINAL PAPER

Anthony J. Hildreth · Maurice Place

The ability of teacher reports to identify child sexual abuse victims

Received: 26 February 1996 / Accepted: 4 September 1996

Abstract This study concerns the validation of an instrument, the SAB2, designed to measure child behaviour and to help identify emotional problems which arise because the child has been a victim of child sexual abuse. Clinically disturbed children, mainstream school children and sexually abused youngsters were assessed. Reliability and validity tests were carried out and a factor analysis was performed which produced seven factors. Finally, the mean factor scores of 35 sexually abused girls were compared with a control sample of 112 girls. The sexual abuse group was found to have significantly higher mean scores than controls on all factors and on the total instrument score. However, closer examination of the scores revealed that 15 of the female sexually abused group did not display any problematic behaviours, whereas the other 20 presented with high scores on all factors and a significantly higher score than any other group on the sexual behaviour factor. The instrument was capable of distinguishing behavioural disturbance arising from sexual abuse from that arising from other sources. However, the instrument failed to detect the 15 'silent' abused cases. It is unclear whether this sub-group has protective factors operating which minimise (or help to mask) the impact of sexual abuse, or whether the effects of the abuse will manifest later in life.

Key words Child sexual abuse · Child behaviour · Teacher reports

Introduction

The past 20 years have seen a rapid increase in clinical interest about the victims of child sexual abuse. This research interest has been of a most diverse nature and has helped us to refine our understanding both of the mecha-

nisms which children use to cope with abuse (Leitenberg et al. 1992) and of the various symptoms which the children may demonstrate. In his review Finkelhor (1988) concluded that the most common symptoms seen in clinical samples are fears, aggressiveness and inappropriate sexual behaviour. However, it is well recognised that there is a great diversity of symptoms, with youngsters demonstrating features such as anxiety and withdrawal (Adams-Tucker 1982; Pascoe and Duterte 1981), conversion disorders (Lewis and Sarrel 1969) and phobic disorders (Peters 1976; MacVickar 1979).

With such a diversity of presenting symptoms, the professional's ability to identify that a child is expressing distress at having been the victim of child sexual abuse depends on a high level of suspicion and an ability to obtain historical details which give the clinical suspicion substance.

In all branches of enquiry the ability to develop uniform methods of identification and evaluation is a major milestone in improving both recognition and intervention. The ability to develop a screening instrument which can, with reasonable accuracy, identify children who would require more detailed enquiry would obviously assist in this process (Snaith 1981).

For such purposes it is usual to seek an instrument which is fairly brief and is both reliable and valid, and many researchers have used schools as the logical setting for sampling the majority of the child population (e.g. Rutter et al. 1970; Kolvin et al. 1981). Although the observations being sought in psychological studies are usually quite involved, there is ample evidence, as Lewine and colleagues (1978) observed, that "teachers are capable of making sophisticated and valuable observations regarding the psychological adjustment of their pupils".

The Social Adjustment and Behaviour 2 (SAB 2) instrument was conceived as a school-based instrument which would attempt to address the question of whether or not sexually abused youngsters present problem behaviours which differ observably from those exhibited by other youngsters who have experienced other trauma-related psychopathology or emotional disturbance.

A. J. Hildreth (✉) · M. Place
Department of Child, Adolescent, and Family Psychiatry,
Sunderland District General Hospital, Kayll Road,
Sunderland, SR4 7TP, England

Subjects and methods

The subjects of this study were 35 girls and 6 boys who had been referred to the Social Services via the Child Protection Process, 73 boys and 6 girls attending local schools and residential centres for children exhibiting emotional problems and disturbed behaviour, and 203 control children from local schools. The study examined three sample groups in order to assess the instrument's performance over as wide a range of behaviours as possible and to allow comparisons to be made. A total of 323 children took part in the project.

Sexually abused group

Criteria for inclusion in the sexually abused group were: girls and boys aged 5–16 years who had been the victims of at least one serious sexual assault, the assault had occurred in the past 12 months and the child was currently on the Child Protection Register. All cases of sexual abuse were notified to the researchers at the time of referral whereupon the school was contacted and the instrument was completed in 4 weeks, school holidays permitting. Each case was subsequently reviewed with the Custodian of the Child Protection Register to assess the degree, nature and duration of the assault(s). Unconfirmed cases were excluded from the analysis.

Disturbed group

Criteria for inclusion into the disturbed group were: boys and girls aged 5–16 years who had undergone a process of psychological assessment under section V of the Education Act 1989 and had been designated as being in need of specialised treatment because of behavioural difficulties. The children either attended a special day school within the area, or were accommodated in a residential establishment where they received in-house schooling.

Control group

The control group consisted of 203 girls and boys aged 5–16 years chosen at random and attending local primary and secondary schools.

The mean ages for the groups were as follows: entire sample 10.7 (SD 3.37), female controls 10.2, female disturbed 12.3, female abused 11.7, male control 10.0, male disturbed 12.2 and male abused 9.3. Table 1 shows the age and gender distribution of the entire sample.

Assessment method

A new 'hybrid' instrument, the Social Adjustment and Behaviour 2 (SAB2), was completed by class teachers on all of the youngsters taking part in the study.

The SAB2 questionnaire was designed as a 50-item measure for use with children in the 5–16 years age range. (The SAB1, which was designed for use with pre-school children, will be the subject of a separate paper.) The questionnaire was intended to serve both as a general screening instrument and as a tool for use with child sexual abuse investigations. The instrument draws on the De-

vereux Child Behaviour Rating Scale (Spivack and Spotts 1966), the Devereux Adolescent Behaviour Scale (Spivack et al. 1967) and the Bristol Social and Adjustment Scale (Stott 1974) for many of the statements (with other items being added to identify behaviours which the clinical literature has indicated are often evident in victims of child sexual abuse). The instrument appears in Table 2.

Items 1–26 are intended to identify the frequencies of certain behaviours. For example, item 1 ('take things that do not belong to her/him'), and item 5 ('display sexuality in play, speech or drawings'), are quantitative elements and are scored from 1 (never) to 5 (very frequently). Items 27–48 are concerned with the degree of a child's propensity to engage in or withdraw from certain actions and situations, for example, item 44 ['preoccupied with compulsive acts s/he recognises as unreasonable, but cannot stop doing (e.g. touching, counting certain acts or routines, etc.)'] and item 33 ('afraid of getting hurt in physical activities'). These items are scored from 1 (not at all) to 7 (extremely). At the end of the questionnaire there are two more specific questions. Item 49 is in two parts and asks whether: (a) 'the child ever exhibited any hostile action or act of symbolic destruction towards father' and (b) '... towards mother'. This is an act which it has been suggested may be symptomatic of abuse (Baker 1983). The last question (not numbered) asks whether the child is known to suffer any illnesses which may be exacerbated by stressful events such as asthma, ulcerative colitis, peptic ulceration and psoriasis.

Statistical tests

Inter-rater reliability and test re-test reliability were computed using Pearson's correlation and Cohen's kappa. Cronbach's alpha was used to test internal consistency, and the reliability coefficient for the whole test was estimated using the Spearman-Brown formula. Specificity and sensitivity were also estimated in order to test the instrument's ability to detect positive and negative cases. After initial examination of the correlation matrix the item scores were subjected to factor analysis using the 'Factor' option in SPSS. Initial factors were extracted by the principal-component method and subjected to Varimax rotation. Hierarchical cluster analysis was also used in the analysis, based on the average linkage between groups.

Mean factor scores for the sexual abuse group and the female control sample were compared using the *t*-test for independent samples.

Results

Inter-rater reliability

A total of 20 teachers took part in the inter-rater testing, all of whom were provided with type-written instructions. Thirty children attending schools for emotional and behavioural problems and 20 mainstream school children were each rated independently by two of their teachers. The age of the sample was an even spread from 5 to 16 years. Using the Pearson Product-Moment Correlation coefficient, the inter-rater reliability of the SAB2 total (ac-

Table 1 Total study population by age and gender

Age (year)	5	6	7	8	9	10	11	12	13	14	15	16	Row total	Distribution of genders
Sexual abuse	3	6	0	1	3	2	3	4	2	7	6	4	41	35 F, 6 M
Disturbed	1	2	2	3	5	6	6	10	15	16	9	4	79	6 F, 73 M
Control	23	23	9	18	24	8	20	20	18	18	12	10	203	112 F, 91 M
Column total	27	31	11	22	32	16	29	34	35	41	27	18	323	153 F, 170 M

Table 2 SAB 2 Instrument – children aged 5–16 years

Child's Name..... School.....

Date of Birth..... Date of Record.....

Teacher..... How Long Known Child?.....

You are going to rate the behaviour of a pupil, for items 1 to 26, use the rating scale below. Please record your score for each item in the box to the right of the statement.

Very Frequently Often Occasionally Rarely Never
5 4 3 2 1

Compared to "normal" children of her or his age, HOW OFTEN does the child:

	ITEM	RATING
1.	Take things that do not belong to her/him?	
2.	Resist, or refuse to do what is asked of her/him, or display a negative attitude?	
3.	Get easily upset by peers when teased, pushed etc.?	
4.	Express the belief that s/he has committed some unpardonable act, that s/he is evil, or that s/he deserves severe punishment?	
5.	Display sexuality in play, speech or drawings?	
6.	Display odd facial grimaces, strange postures, or odd movements, e.g. hitting or biting her/himself, senseless or magical movements of the fingers, arms, legs, or head etc.?	
7.	Show an interest in violence, death, people in accidents (e.g. in what s/he reads, talks about, watches on television, etc. ?	
8.	Tend to be loud and boisterous?	
9.	Have a fixed facial expression that lacks feeling?	
10.	Tend to cling to adults (e.g. want to sit next to, be around them a lot, etc.?	
11.	Do what s/he wants to do, even when told not to?	
12.	Get very upset or emotional if things don't go her/his way?	
13.	Seek out adult approval for what s/he has done?	
14.	Appear overactive and constantly moving about	
15.	Mumble, shout out, or make unusual vocal noises?	
16.	Make sexual advances to adults?	
17.	Express depressed or despairing thoughts (e.g. express lack of hope, feelings of discouragement, that s/he expects the worst, no sense in trying etc. ?	
18.	Persist when told s/he can't have something (e.g. nag, demand, repeatedly ask for something) ?	
19.	Exhibit an interest in sex through action, or what s/he says?	
20.	Express strange grandiose ideas about her/himself (e.g. that s/he has an unusual or fantastic power over others or things, that s/he is an extremely important person, etc.) ?	
21.	Act bossy with other children?	
22.	Blame or condemn self for things that happen to her/him?	
23.	Acts before s/he thinks (i.e. is impulsive) ?	
24.	Show a reluctance to go home from school	
25.	Dress in pseudo-mature clothing? (younger child) or sexually provocative clothing (older child)?	
26.	Stay off school with reported minor illnesses?	

Table 2 (continued)

FOR ITEMS 27 to 48 PLEASE USE THE RATING SCALE BELOW:

Extremely 7 Distinctly 6 Quite a bit 5 Moderately 4 A Little 3 Very slightly 2 Not at all 1

To what degree is the child:

	ITEM	RATING
27.	Socially inhibited?	
28.	Prone to hit or Physically threaten peers?	
29.	Boycrazy (for girls) Girlcrazy (for boys)?	
30.	Indifferent to new learning tasks?	
31.	Unemotional, rarely shows feelings?	
32.	Obsessed or preoccupied with ideas s/he worries or talks a lot about?	
33.	Afraid of getting hurt in physical activities?	
34.	Changeable or variable in mood or emotional state?	
35.	Prone to tire quickly or have low endurance?	
36.	Sneaky or underhand in much of what s/he does?	
37.	Unpopular with peers?	
38.	Prone to truancy or running away?	
39.	Promiscuous?	
40.	Prone to self mutilation (e.g. burning, self-poisoning, cutting/scratching skin)?	
41.	Too lacking in energy to ask the teacher's help?	
42.	Preoccupied with cosmetics (e.g. eye shadow, rouge, after shave, hair lotion etc.)?	
43.	One whose contacts with the opposite sex must be supervised?	
44.	Preoccupied with compulsive acts s/he recognises as unreasonable, but can't stop doing (e.g. touching, counting certain acts or routines etc.)?	
45.	Prone to keep her/his distance from adults?	
46.	Obsessed with cleanliness? (cleans self)	
47.	Failing to meet expected academic performance?	
48.	Unmotivated or lacking in sufficient energy to get on with work by her/himself?	
49.	Has the child ever exhibited any hostile action or act of symbolic destruction towards one or both parents?	tick box if applies
(i)	Towards father?	
(ii)	Towards mother?	
	Where the answer is "yes", please describe the act(s) carried out and rate HOW OFTEN using the rating scale (1 to 5) from the first section:	RATING
	Towards father	
	Towards mother	

If the child is known to suffer from any ailment such as asthma, psoriasis, ulcerative colitis, or peptic ulceration, which may be exacerbated by fluctuations of the mental state, please describe

Are there any comments you wish to make regarding this child?

Are you concerned about this child's behaviour? No [] Possibly [] Yes [] (tick)

Thank you for completing this form

cumulated) score was 0.87 ($P < 0.0001$, $n = 50$). The inter-rater reliability of all 49 individual items was examined and ranged from 0.43 for 'seeking out adult approval for what s/he does', to 0.90 for 'prone to self mutilation'. The average correlation for all the items was 0.75. All items were significant to $P < 0.0001$, except item 13 'seeks adult approval', which was significant to $P < 0.002$, and the item relating to symbolic destruction of parents, which was not significant.

One possible concern with correlation scores is that a high correlation coefficient may arise between raters if they differ systematically. However, the simple precaution of examining rating columns for any such uniformity allows to determine whether or not the correlation score does, in fact, represent a true level of congruity. In this case it was apparent that the correlation scores were indicative of agreement between individual teachers.

Cohen's kappa

Inter-rater reliability can also be assessed using a version of the kappa coefficient which was developed by Cohen (1960) to assess reliability of diagnosis. In this case it was used simply to test the rate of agreement achieved by teachers as to whether or not a child's total score on the instrument was over or under a cutoff score. With a cutoff score of 100 the kappa for rate of agreement was 0.71 ($n = 49$) with an observed rate of agreement of 86%. Although it may appear to be a convenient figure, the value of 100 for a cutoff point was not chosen because of its numerical convenience, rather that the sensitivity vs 1-specificity tradeoff appeared optimum at this score.

Test re-test reliability

A total of 75 children made up of 41 disturbed pupils from special schools and 34 pupils from mainstream schools were assessed at the beginning and end of a 6 week period in order to examine the instrument's ability to return the same results over a period of time. The test re-test reliability for the instrument total score was 0.91 ($P < 0.0001$). Correlation scores for 48 of the individual items ranged from 0.89 for 'exhibit an interest in sex, through action or what s/he says', to 0.41 for 'blame or condemn self for things that happen to him/her' (all significant to $P < 0.0001$). Only item 49, relating to symbolic destruction of parents, was not significant. The average correlation of all items was 0.75.

Internal consistency

The reliability of the SAB2 was tested using Cronbach's alpha to examine its internal consistency. Alpha for the 50-item scale was 0.956 ($n = 323$), which indicates that the scale is reliable, even allowing for the fact that the scale is fairly long (Norris 1992).

Split-half reliability

A split-half reliability analysis was performed, and it was found that the correlation between the two halves of the form was 0.78, the Spearman-Brown coefficient (equal length) was 0.88, the Cronbach's alpha coefficient for the first half was 0.94 and for the second half was 0.91. The reliability of the overall test can also be estimated using the Guttman split-half coefficient. Unlike the Spearman-Brown test, this test does not assume that the two parts are equally reliable or have the same variance. The instrument was found to have a Guttman split-half reliability of 0.88 which is reasonable for a measure of this length. The Guttman test is seen as more rigorous than the Spearman-Brown (Norris 1992).

Specificity and sensitivity

An instrument's discriminative power may be measured by its ability to identify true cases (sensitivity) and by the rate at which non-pathological cases are identified correctly (specificity). Problems may be encountered where a questionnaire either fails to detect pathology because the instrument is insufficiently sensitive to abnormal behaviour (lacks sensitivity), or where the instrument wrongly indicates a problem where none exists (lacks specificity). Several cutoff scores were investigated to determine maximum efficiency, and it was found (above) that a cutoff score of 100 correlated well with external assessments of disturbance. The performance of the instrument was then examined for each of the two study groups, disturbed children and sexually abused children, in order to evaluate its ability to detect pathology. Firstly, children who had been classified as 'clinically disturbed' by a process of psychiatric assessment were coded. This was used to determine the instrument's ability to detect pathology. Any children who were known to have been sexually abused were excluded from this analysis. The data were then analysed to determine the sensitivity, specificity, and overall efficiency of the instrument. Using these groups (disturbed vs controls, $n = 282$) the instrument was found to have a sensitivity of 83%, a specificity of 86% and an overall efficiency of 85%. However, these values may be exaggerated somewhat as the cutoff score is derived wholly from the sample data.

Secondly, the sexually abused group of children were compared with the control sample and the tests were re-applied. Using these groups ($n = 244$), sensitivity was 61% and specificity was 86% with overall efficiency at 82%. The relatively low sensitivity can be largely explained by the fact that 20 of the female sexually abused group (57%) displayed clear behavioural problems, whereas 15 (43%) showed no detectable pathology.

Concurrent validity

To further test the validity of the instrument, the SAB 2 total scores were compared with scores obtained on the

Rutter B instrument which had been completed on the same occasion. A Pearson correlation of 0.91 was obtained ($n = 250$) indicating a high level of concurrent validity.

Factor analysis

Having tested the instrument's ability to discriminate disturbed and sexually abused cases from control cases, the

issue of the instrument's capacity to identify sexual abuse victims was addressed further.

The aim of factor analysis is to describe the covariance relationship among a number of variables in terms of a few underlying random quantities called factors (Johnson and Wichern 1992). Put more simply, if variables are grouped according to the magnitude of their correlations with one another so that all variables in a particular group are highly correlated, but have relatively small correlations with variables in other groups, it is possible that a

Table 3 Factors

Items	1. Over-act	2. Sexual	3. Moti-vation	4. Inse-cure	5. With-drawn	6. Nega-tive	7. Dis-turbed
11. Dowhat	0.841 C						
2. Resists	0.798 C						
18. Persists	0.766 C						
28. Hits	0.750 C						
8. Loud	0.717 C						
13. Impulsive	0.709 C						
1. Takes	0.700 C						
21. Bosses	0.677 C						
36. Sneaky	0.663 C						
12. Emotion	0.648			0.473 C			
34. Moody	0.615			0.412 C			
14. Overact	0.508 C						
3. Easily upset	0.508			0.408 C			
43. Oppsex		0.842 C					
39. Promiscuous		0.792 C					
19. Sex interest		0.789 C					
29. Boy/girl crazy		0.773 C					
5. Sexual play		0.726 C					
16. Advances		0.699 C					
25. Dress		0.461 C					
42. Cosmetic		0.453 C					
40. Selfharm		0.334*C					
48. Unmotivated			0.844 C				
41. Energy			0.762 C				
47. Academic			0.730 C				
30. Nolearn			0.711 C				
35. Tires			0.634 C				
13. Approval				0.792 C			
10. Clings				0.763 C			
32. Worries				0.611		0.323 C*	
33. Nohurt			0.344*C	0.473			
31. Unemotional					0.774 C		
9. Fixed					0.762 C		
27. Inhibited					0.550 C		
45. Distance					0.541 C		
37. Peers					0.514 C		
22. Blames						0.857 C	
4. Evilact						0.778 C	
17. Despairs						0.640 C	
24. Nohome						0.429 C	
6. Tics							0.733 C
15. Mumbles							0.578
44. OCD							0.472 C
7. Int. Violence							0.452 C
20. VIP							0.366*C

single underlying construct is responsible for the observed correlations. We chose to analyse the data using this method because it was felt that if certain items tended to band together into the types of behaviours suggested by researchers as being collectively indicative of abuse (see e.g. Finkelhor 1988), then differences in factor scores achieved between abused and non-abused children may be suggestive of sexual abuse.

The sample chosen for the factor analysis consisted of 175 children. A total of 79 youngsters were included who had been identified by either an educational psychologist or a child psychiatrist as having a behavioural/emotional problem. This group included children at special educational establishments, and children referred via the child protection procedure whose cases did not satisfy the criteria for the sexual abuse group. A total of 41 children who were victims of sexual abuse and 55 other youngsters (attending mainstream schools) about whom teachers had expressed concern over their behaviour were also included. Sampling sufficiency was assessed using the Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy. This test is used to test the appropriateness of the factor model. A measure of 0.82 was obtained which, according to Kaiser (1974) 'meritorious' indicating that the factor model is highly appropriate. Bartlett's Test of Sphericity was used to test the hypothesis that the correlation matrix is an identity matrix. A low significance level indicates that the relationship between the variables is weak and therefore the factor model is inappropriate. A score of

5860.2 ($P < 0.000001$) was attained indicating that a strong relationship exists between the variables.

The analysis produced seven factors from 45 of the 49 scoreable items in the SAB2. The four items which did not group with other factors were those relating to symbolic destruction of parents (composite item 49), items concerning staying off school with minor illnesses' and 'truanting' (items 26 and 38), and item 46 regarding 'obsessive cleaning'. Table 3 shows the factor loadings and the equivalent cluster groupings ('C' denotes items which were grouped together when a cluster analysis was performed).

Factor scores

The factor scores for the female sexual abuse sample ($n = 35$) were computed and compared against those for the female control sample ($n = 112$) and it was found that for all factors the sexual abuse group had a significantly higher mean score. (Table 4).

On further examination, however, the standard deviations for the sexual abuse group pointed to a large variation in scoring within the group. The overall average score disguised the fact that 15 of the group scored 'normally', whereas 20 youngsters had scores indicating a very high level of disturbance. The variation in the scoring is shown in Table 5.

Finally, a comparison was made between the factor scores of all groups studied (Table 6). It can be seen that

Table 4 SAB2 factor scores for sexual abuse and control groups (females)

Factor name	Max. score	Female sexual abuse ($n = 35$)		Female control sample ($n = 112$)		<i>t</i> -test significance	95% CI for differences
		Mean score	SD	Mean score	SD		
Overact	71	30.3	14.7	20.4	8.9	$P = 0.001$	5.8–13.9
Sexual	55	14.1	8.5	10.2	3.1	$P = 0.013$	2.0– 5.7
Motivation	35	14.9	9.2	7.3	3.4	$P < 0.0001$	5.6– 9.6
Insecure	24	9.4	4.7	7.1	2.7	$P = 0.009$	1.1– 3.6
Withdrawn	33	14.5	7.1	9.4	3.8	$P = 0.0001$	3.2– 6.9
Negative	20	7.0	4.0	4.4	1.0	$P = 0.001$	1.7– 3.4
Disturbed	27	8.4	4.2	5.5	1.2	$P = 0.001$	2.0– 3.8
Total SAB 2 score	294	106.7	39.0	70.4	19.2	$P < 0.0001$	26.7–46.0

NOTE: Confidence interval calculations based on Levene's test for equality/inequality of variances. $P < 0.0001$ in all cases

Table 5 Variation in scoring of sexually abused sample (females)

Factor name	Mean score: girls acting out group ($n = 20$)		Mean score: no symptoms ($n = 15$)		<i>t</i> -test significance	95% CI for differences
	Mean score	SD	Mean score	SD		
Overact	38.7	13.5	19.2	6.9	$P < 0.0001$	11.7–27.3
Sexual	17.3	10.1	9.8	1.8	$P = 0.008$	2.1–12.9
Motivation	19.9	7.7	8.3	6.4	$P < 0.0001$	6.8–16.5
Insecure	11.2	5.2	7.1	2.8	$P = 0.006$	1.3– 6.9
Withdrawn	17.5	6.8	10.5	5.5	$P = 0.002$	2.6–11.2
Negative	9.0	4.3	4.3	0.7	$P < 0.0001$	2.3– 6.9
Disturbed	11.0	4.1	5.1	0.3	$P < 0.0001$	3.8– 8.0
Total SAB 2 score	134.0	27.2	70.3	29.3	$P < 0.0001$	48.0–79.4

NOTE: Confidence interval calculations based on Levene's test for equality/inequality of variances

Table 6 Factor scores for all groups

Factor name (maximum score in brackets)	Female control (<i>n</i> = 112)		Female dis- turbed (<i>n</i> = 6)		Femal abuse (no symptoms; <i>n</i> = 15)		Female abuse (acting out; <i>n</i> = 20)		Male control (<i>n</i> = 91)		Male disturbed (<i>n</i> = 73)		Male abus (<i>n</i> = 6)	
	Mean score	SD	Mean score	SD	Mean score	SD	Mean score	SD	Mean score	SD	Mean score	SD	Mean score	SD
(71) Overact	20.4	8.9	25.2	6.7	19.2	6.9	38.7	13.5	24.9	9.7	43.6	13.2	40.3	11.1
(55) Sexual	10.2	3.1	12.0	3.7	9.8	1.8	17.3	10.1	10.1	2.6	13.6	7.3	12.7	2.2
(35) Motivation	7.3	3.4	16.8	6.1	8.3	6.4	19.9	7.7	10.0	6.5	17.1	7.3	15.8	5.1
(24) Insecure	7.1	2.7	12.7	7.0	7.1	2.8	11.2	5.2	6.8	2.9	11.0	5.0	14.7	4.6
(33) Withdrawn	9.4	3.8	20.2	8.3	10.5	5.5	17.5	6.8	10.3	4.5	15.4	5.9	15.2	6.0
(20) Negative	4.4	1.0	7.2	3.2	4.3	0.7	9.0	4.3	4.6	1.1	7.8	3.5	10.7	4.5
(27) Disturbed	5.5	1.2	6.8	2.6	5.1	0.3	11.0	4.1	6.7	2.3	11.4	5.9	9.7	6.7

the female abuse group who do not act out have similar scores to the female control group. The female 'acting out group', on the other hand, score highest on factor 2 (sexual), factor 3 (lacking motivation) and factor 5 (withdrawn). This group also scores high on the other four factors along with the male disturbed and male sexual abuse groups. The male abuse group (*n* = 6) had the highest score on factor 6 (negative) while also scoring highly (second place) on factor 1 (overact). They also scored highly on all the other factors, achieving either second or third place.

Discussion

The Questionnaire Model: appropriateness to child sexual abuse

To develop an instrument for detecting behavioural signs of child sexual abuse within the classroom setting, there must be a tacit acceptance that some child victims display observable manifestations of abuse, and many authors have published research findings to support this hypothesis (see e.g. the work by Finkelhor 1988; Adams-Tucker 1982; Pascoe and Duterte 1981; Peters 1976; Lewis and Sarrel 1969; Tufts' 1984). Secondly, there must be the assumption that classroom-based measurement of child behaviour is an appropriate model. Again, several authorities support this view (see e.g. Miller 1967; Rutter 1967; Connors 1969; Edelbrock and Achenbach 1984). Therefore, given that behavioural signs of child sexual abuse may be discernible, and that emotional and behavioural problems may be measured in the classroom, it was felt that the concept behind the project was a valid one. The whole credibility of the project, however, rested on the hypothesis that teachers would also be able to detect the types of behavioural change frequently associated with sexual abuse.

The other factor to consider when deciding how to assess emotional and behavioural problems in children is whether to use self-report or other-person completion. Because of the nature of the investigations with which this project was concerned, it was considered that self-assess-

ment methods were inappropriate for a number of reasons. Firstly, because of the age-range of the subjects, it could not be assumed that they would all be able to read and understand the questionnaires. Secondly, a high degree of detachment and objectivity is required to accurately assess one's own condition and, given the nature of the trauma experienced, it was felt that any responses could not be properly relied upon. Finally, one vital difference exists in the content of the two types of scale: other-person rating scales examine signs *and* symptoms, whereas self-report, by its very nature, is restricted to symptoms alone (Zerssen 1980). Considering all these factors it seemed most appropriate to use an other-person instrument which utilised information from someone who knew the child well and had a sufficiently detached and unbiased viewpoint. This set of requirements makes the teacher ideally placed to act as a source of screening information.

In any questionnaire it is vital that ambiguity is eliminated as much as possible. Helzer et al. (1977) pointed out that the raters' understanding of the questionnaire must coincide with that of the researcher if the findings are to have any meaning. In order to attempt to ensure an understanding of what was required of them, teachers were given written instructions with each set of questionnaires and given a verbal presentation regarding completion of the forms and the reasons for the research.

A final consideration with regard to questionnaire methods of assessment is the distribution of the target symptom in the sample population. Goldberg (1972) pointed out that if the estimate of an individual's disturbance based on a score on a screening device is to be an accurate assessment of the degree of that person's psychological disturbance, the level of that psychiatric disturbance in the general population should have a normal distribution. In other words, it must be assumed that psychiatric disturbance is evenly distributed throughout the population in varying degrees of intensity. Once this is established, it is possible to allocate a cutoff score above which disturbance may be inferred. Much, therefore, depends on the validation of the instruments in question and the validity of the research methods employed.

Reliability

The present instrument offers a pattern of reliability and validity which suggests that it is worthy of further use. The SAB2 instrument appears able to identify sexually abused children who 'act out', and to distinguish such youngsters from other 'disturbed' children. There is, however, a clear difference in the scoring for one factor, factor 2 (sexual). Here the female 'acting out' abuse group scored much higher than all other groups and was significantly higher than the next highest group, the male disturbed (t -value = 1.82, $P < 0.05$ with 91 *df*). The 'non-acting out' female abuse group scored lowest on this factor which suggests that sexually abused children (girls at least) fall into two distinct groups, those who display problem behaviours and those who do not. The reason that such differences may exist is not clear. From our sample we noted that age does not appear to play a part, nor does socio-economic factors. Our data did not permit an analysis of the amount of support available to the child within the family, nor was there data pertaining to whether the handling of individual cases affected the outcome, as has been suggested by Gomes-Schwartz et al. (1990).

The difference could also arise because the type of sexual abuse has prompted a different degree of response from professionals. Some studies have reported that father-child incest has the worst effects upon the child (Adams-Tucker 1982; McLeer et al. 1988), whereas Finkelhor (1979) argues that there is no clear difference between the effects of abuse perpetrated by immediate family members, other relations, friends and strangers. All of the youngsters in this study sample were abused by a close family or substitute family member. Similar controversy exists as to whether the duration of the abuse is significant to the outcome (see e.g. Gagnon 1965, and Finkelhor 1979).

One further possibility is that some children have resources for coping with adversity which protects them from developing marked disturbance of emotions or behaviour (Rutter 1985). Such protective factors fall into three broad categories (Garmezy et al. 1984): factors within the child, in the family and in the wider community. Unfortunately, the scope of this study did not permit us to explore the contribution which such factors may be making to this sub-group's relative well-being.

Conclusions

This study suggests that the SAB 2 is a valid and reliable instrument for school-based screening of children who are showing an 'acting-out' pattern of behaviour in response to being the victim of sexual abuse. The instrument does, however, fail to detect the 'silent' abused cases where the child shows little disturbance or behavioural reaction. Whether the results indicate that the instrument is not sufficiently able to identify behaviours resulting from sexual abuse, or whether these youngsters do, as suggested, represent a significant sub-group who have protective factors

operating which minimise (or help to mask) the impact of abuse on them, will require further work to clarify. Finally, it must be stressed that the intention was not to develop a questionnaire capable on its own of identifying youngsters who have been sexually abused, but rather to assist in the identification of problems arising from sexual abuse. Although such problems are broadly similar to those seen in children with a variety of traumatic experiences, the child who has been the victim of sexual abuse does appear to demonstrate a pattern which is somewhat specific in makeup. If confirmed, this may allow more focused interventions to be developed which can be used at the early stages of detection to help minimise the evolution of these difficulties.

References

- Adams-Tucker C (1982) Proximate effects of sexual abuse in childhood: a report on 28 children. *Am J Psychiatry* 13: 1252-1256
- Baker T (1983) Report on reader survey: child sexual abuse: "19" confidential survey (mimeo). St. George's Hospital, London
- Cohen J (1960) A coefficient of agreement for nominal scales. *Educ Psychol Measure* 29: 37-46
- Conners CK (1969) A teacher rating scale for use in drug studies with children. *Am J Psychiatry* 12: 884-888
- Edelbrock C, Achenbach TM (1984) The teacher version of the Child Behaviour Profile: I boys aged 6-11. *J Abnorm Child Psychol* 52: 207-217
- Finkelhor D (1979) Sexually victimised children. Free Press, New York
- Finkelhor D (1988) The trauma of child sexual abuse: two models. In: Wyatt GE, Powell GJ (eds) *Lasting effects of child sexual abuse*. Sage, Newberry Park
- Gagnon J (1965) Female child victims of sex offenses. *Soc Problems* 13: 176-192
- Garmezy N, Masten AS, Tellegen A (1984) The study of stress and competence in children: a building block for developmental psychopathology. *Child Dev* 55 (1): 97-111
- Goldberg DP (1972) The detection of psychiatric illness by questionnaire. Oxford University Press, London
- Gomes-Schwartz B, Horowitz J, Cardarelli AP, Sauzier M (1990) The aftermath of child sexual abuse: 18 months later. In: Gomes-Schwartz B, Horowitz J, Cardarelli AP (eds) *Child sexual abuse: the initial effects*. Sage, Newberry Park
- Helzer JE, Robins LN, Tableson M, Woodruff RA, Reich T, Wish GD (1977) Reliability of psychiatric diagnosis. *Arch Gen Psychiatry* 34: 129-133
- Johnson RA, Wichern DW (1992) *Applied multivariate analysis*. Englewood Cliffs, Prentice-Hall, New Jersey
- Kaiser HF (1974) An index of factorial simplicity. *Psychometrika* 35: 401-415. Cited in Norusis MJ (1990) *SPSS Statistics 4.0*. SPSS Inc., Chicago
- Kolvin I, Garside RF, Nicol AR, MacMillan A, Wolstenholm F, Leitch IM (1981) *Help starts here: the maladjusted child in the ordinary school*. Tavistock Publications, London
- Leitenburg H, Greenwald E, Cado S (1992) A retrospective study of long-term methods of coping with having been sexually abused during childhood. *Child Abuse Neglect* 16(3): 399-407
- Lewine RRJ, Watt NF, Prentky RA, Fryer JH (1978) Childhood behaviour in schizophrenia, personality disorder, depression and neurosis. *Br J Psychiatry* 132: 347-357
- Lewis M, Sarrel PM (1969) Some psychological aspects of seduction, incest, and rape in childhood. *J Am Acad Child Psychiatry* 8: 606-619
- MacVicar K (1979) Psychotherapeutic issues in the treatment of sexually abused girls. *J Am Acad Child Psychiatry* 19: 342-353

- McGuire J, Richman N (1988) Pre-School Behaviour Checklist Handbook. NFER-NELSON, Windsor
- McLeer SV, Beblinger E, Atkins MS, Foa EB, Ralphe DL (1988) Post-traumatic stress disorder in sexually abused children. *J Am Acad Child Adolesc Psychiatry* 27: 650–654
- Miller LC (1967) Louisville Behaviour Checklist for males, 6–12 years of age. *Psychol Rep* 21: 885–896
- Norusis MJ (1990) SPSS statistics 4.0. SPSS Inc., Chicago
- Norusis MJ (1992) SPSS for Windows: professional statistics, release 5. SPSS Inc., Chicago
- Pascoe DJ, Duterte BO (1981) The medical diagnosis of sexual abuse in the premenarchal child. *Pediatr Ann* 10 (5): 40–45
- Peters J (1976) Children who were victims of sexual assault and the psychology of offenders. *Am J Psychother* 30: 398–421
- Rutter M (1967) A children's behaviour questionnaire for completion by teachers: preliminary findings. *J Child Psychol Psychiatry* 8: 1–11
- Rutter M (1985) The developmental psychopathology of depression: issues and perspectives. In: Rutter M, Izard C, Read P (eds) *Depression in childhood: developmental perspectives*. Guilford Press, New York
- Rutter M, Tizard J, Whitmore K (eds) (1970) *Education health and behaviour*. Longmans, London
- Snaith RP (1981) Rating scales. *Br J Psychiatry* 138: 512–514
- Spivack G, Spotts J (1966) The Devereux child behaviour rating scale: manual. The Devereux Foundation Press, Devon
- Spivack G, Haimes PE, Spotts J (1967) Devereux adolescent rating scale. The Devereux Foundation, Pittsburg
- Stott DH (1974) The social adjustment of children. Manual of the Bristol Social Adjustment Guides, 5th edn. Hodder and Stoughton, London
- Tufts' New England Medical Centre, Division of Child Psychiatry (1984) Sexually exploited children: service and research project (Final report for the Office of Juvenile Justice and Delinquency Prevention). U.S. Department of Justice, Washington, DC
- Zerssen D von (1980) Psychopathomethrische Verfahren und ihre Anwendung in der Psychiatrie. In: Peters UH (ed) *Die Psychologie des 20. Jahrhunderts*. Kindler, Zürich, pp 149–169