British Journal of Psychiatry (1989), **155**, 24–35

Adult Personality Functioning Assessment (APFA)
An Investigator-Based Standardised Interview

J. HILL, R. HARRINGTON, H. FUDGE, M. RUTTER, and A. PICKLES

The development of an investigator-based standardised interview to assess patterns of specific and general social dysfunction is described. It covers six domains of functioning: work; love relationships; friendships; non-intimate social contacts; negotiations; and everyday coping. Inter-rater reliability was tested by three investigators rating 21 audiotaped interviews, and was shown to be high, with an intraclass correlation of 0.87 for the total score. The pattern of associations between specific and general social dysfunction was examined through determination of sensitivities and specificities and through LISREL modelling. The findings varied across social domains but it was concluded that the total APFA score provided a reasonable measure of general social dysfunction.

During recent years there has been the development of a range of standardised questionnaire and interview measures of personality disorder as conceptualised in both DSM–III and ICD–9 (Merikangas & Weissman, 1986; Reich, 1985; Tyrer & Ferguson, 1987). In contrast to the poor agreement on the diagnosis of personality disorder generally found between clinicians using non-standardised interviews, several of these standardised techniques have been shown to have fairly good reliability. However, little is known of their validity, and they are limited in being tied to current concepts of personality disorder, which are both contradictory and poorly substantiated (Frances & Widiger, 1986; Rutter, 1987*a*). Those based on ICD–9 (World Health Organization, 1978) approach the diagnosis of personality through the assessment of particular traits represented by such features as dysthmia, aggressivity, and obsessionality (Mann *et al*, 1981; Tyrer *et al*, 1979, 1984). Those based on DSM–III (American Psychiatric Association, 1980) to some extent do the same, but the main emphasis is on particular types of behaviour or states of mind thought to characterise specific personality disorders (Siever & Klar, 1986).

Thus, the diagnosis of antisocial personality disorder is based on consistent violation of the rights of others and lack of loyalty in interpersonal relationships; and that of schizotypal personality disorder on cognitive perceptual distortions and social withdrawal. A problem arises, however, from the finding that the majority of individuals who meet the criteria for one type of personality disorder also meet the criteria for at least one other (Clarkin *et al*, 1983; Mellsop *et al*, 1982; Stangl *et al*, 1985). But the disorders do not overlap randomly; some are closely related to one another whereas others stand out as different both in concept and as measured (Clarkin *et al*, 1983; Rutter, 1987*a*). Probably, those that represent lesser variants of specific psychiatric disorders (such as schizotypal and personality disorder and schizophrenia, or cyclothymic/affective personality disorder and affective disorder) need to be differentiated from the rest. The remainder, although separately defined in terms of particular traits or types of behaviour, have in common a pervasive and persistent abnormality in social functioning. However, none of the existing interview assessments of personality disorder focuses on this as a central feature.

Social functioning has been the focus of the 30 or so scales measuring various aspects of social adjustment. These have been systematically reviewed by Weissman and her colleagues (Weissman, 1975; Weissman *et al*, 1981; John & Weissman, 1987), and an additional interview measure of social role performance has been described by Sturt & Wykes (1987). A range of useful measures is available, many of which have been shown to have satisfactory psychometric properties, but none is intended to assess personality functioning (and could not do so without a major reorientation). Rather, most were designed to evaluate patients' community adjustment either while ill or following recovery. Accordingly, the focus is on the present (rather than lifetime) and on the impact of illness, often with measures of satisfaction and subjective feelings. Most of the scales follow a highly structured questionnaire format, with the limitation that many items apply to only some subjects and that scoring procedures do not take into account the person's social context or situation (Katschnig, 1983).

https://doi.org/10.1192/bjp.155.1.24 Published online by Cambridge University Press

Thus, existing measures that might be used to assess personality functioning tend to concentrate on either temperamental traits and supposedly specific psychological abnormality, or on aspects of current social adjustment and satisfaction. Neither group of measures provides an adequate evaluation of objectively defined social functioning in a range of domains. Such an evaluation is required if personality disorder is conceptualised as requiring a pervasive and persistent abnormality in social functioning. However, in the present state of knowledge, it is necessary that the measurement should not be tied to any one classification scheme (Rutter & Pickles, 1989). Not only do we lack data on the links between social malfunction and deviant temperamental traits or specific psychological abnormalities, but also it is not known how commonly individuals have persistent abnormalities in some social domains but not in others. The question of how pervasive social malfunction must be for personality disorders to be diagnosed requires empirical data as much as theory, if it is to be answered satisfactorily.

Moreover, it is clear that many individuals are likely to show malfunction in particular social domains or in particular situations for reasons that are separate from personality disorder. Thus, often, marital problems will be a function of difficulties in that particular dyadic relationship, or there may be a pervasive difficulty in intimate relationships that is a consequence of, say, insecurities in attachment relationships in childhood, but which is not accompanied by any problems in work or everyday coping.

It seems therefore that what is needed, but what has been lacking, is a standardised measure of a person's functioning over time in a range of different social domains, the measurement being done in a way that can deal with varied life circumstances and which can assess functioning in people without, as well as with, acute psychiatric disorder. The Adult Personality Functioning Assessment (APFA), a standardised investigator-based interview, was designed to fill that gap. In this paper we describe its rationale and format and give findings on its reliability and structure as assessed on the subject version of the interview.

The Adult Personality Functioning Assessment

The interview is primarily designed to provide a standardised assessment of a person's functioning in a range of social domains, with the aim of both identifying dysfunction that is specific to particular domains and measuring general social dysfunction that is pervasive across most social domains. Six domains were selected as representative of the variety of social demands likely to be faced by most people regardless of their circumstances. These are: work; love relationships; friendships; non-intimate social contacts; negotiations; and everyday coping. In each domain, the intention is to determine the person's level of functioning regarding role performance in that domain, without reference to functioning in other social domains.

Rating of level and type of functioning

For each domain, the rating of level of functioning is made on a six-point scale extending from '0' (unusually effective in the operative life circumstances to a degree that is clearly above average to a substantial degree) to '5' (pervasive failure of role performance in the specified domain). The latter would apply, for example, in work if the subject had failed to hold any job for more than brief periods; and in love relationships if there were an absence of sustained committed relationships or if marriage/cohabitation were maintained only in the face of open hostility or apathy leading to essentially separate lives.

The process of rating necessarily involves the weighing up of a complex mixture of different aspects of functioning, with the need to balance elements of successful role performance against elements of role failure. Comparability is ensured by four features. First, for each domain there is explicit specification of the elements to be taken into account and how they are to be combined. For example, the rating for 'work' requires attention to conflict with supervisors or fellow-workers, difficulties resulting from poor work performance, frequent changes of job, absences from work, poor time-keeping, and being sacked or walking out from a job. Similarly, the rating for 'non-intimate social contacts' takes into account the frequency, range, and nature of interactions with neighbours and acquaintances and participation in social situations, with particular reference to altercations with people, or avoidance of social participation, or marked circumscribing of topics of conversations with others. These features are summarised in the interview schedule and are outlined in greater detail in the manual that accompanies it (not published here, but both are available at cost price from the authors). Secondly, a 'dictionary' of examples accumulated from the use of the interview in several studies provides details of how various combinations of features should be rated. Thirdly, the schedule provides explicit guidance on the topics to be covered, together with a listing of suggested probes.

https://doi.org/10.1192/bjp.155.1.24 Published online by Cambridge University Press

Thus 'negotiations' requires an account of the subject's performance in commercial and professional transactions with people outside the family. The interviewer is required to ask about fairly routine negotiations such as making purchases in a shop or making an appointment at the doctor or dentist; those involving some explanation and assertion of rights, such as applying for a job or taking an unsatisfactory item back to a shop or discussing a child's progress with teachers; and those that require a point of view to be put with some persistence, as with the failure of a landlord to undertake essential repairs or where there has been a major difference of opinion over some aspect of a child's education. Similarly, everyday coping refers to the responsibility taken for tasks in the home such as cooking meals, washing, looking after children, keeping the house in good order and managing the finances. Unlike the other five domains, this does not involve interpersonal skills.

Fourthly, a step-by-step strategy is followed in rating. The first decision is whether there is substantial dysfunction in the domain; if there is, the rating must be 3, 4 or 5; if not, the rating must be 0, 1 or 2. Ordinarily, any serious dysfunction requires a rating of '3' or worse. Thus, someone who had had temporary marital separations through discord or who had lost several friendships as a result of quarrelling or some kind of rift would rate '3' unless there were some special circumstances. If there is not substantial dysfunction, the next decision is whether the functioning is more than adequate in showing markedly positive features without any significant negative ones. A rating of '0' on love relationships would require not only temporal stability but positive trust, marked confiding and enjoyment. A '1' rating is made where functioning is generally satisfactory and without significant lasting difficulties, but lacking the positive features required for '0'. A '2' rating implies that the level of functioning is mostly satisfactory but there are significant difficulties of one kind or another. The differentiation between 3, 4 and 5 is based on the seriousness of the dysfunction, its extensiveness over situations, and the presence/absence of aspects of good functioning. Thus, a rating of '3' on 'love relationships' might be appropriate if the predominant pattern were one of tension and discord but yet there were some shared activities and a degree of commitment.

For each domain, a separate rating is made of the type of dysfunction. The main differentiation is between dysfunction characterised by arguments, discord, tension, aggression or violence, and that mainly characterised by apathy, lack of involvement or avoidance, but there is also provision for rating 'other' types.

Period of coverage and applicability

People's social circumstances and role performance requirements are likely to vary greatly over the life span and if levels of social performance are to be used in assessing personality functioning, it is necessary to specify which age period is to be used in rating. APFA ordinarily uses the 21–30 year period as a baseline (but provision is made for circumstances in which some other age period has to be used). Three particular considerations must be taken into account in assessing the subject's psychosocial functioning: opportunity, circumstances, and association with illness or psychiatric disorder. Functioning can be assessed only if the person has had the opportunity to function in the specified domain. The domains and age period were chosen to ensure that the great majority of people will have had the necessary opportunities but this will not always be the case (as, for example, when someone has been in prison) and the manual provides guidance on how to deal with this issue. The general principle is that performance should be rated only for time periods when opportunities have been present but that there should be detailed questioning to determine whether the lack of opportunity was a consequence of personal social dysfunction (for example, whether unemployment was a consequence of redundancy or dismissal or failure to seek work). In cases in which ratings on one domain are not applicable through lack of opportunity, scores for the other domains may be prorated to derive a total score. Similar issues arise with respect to circumstances. Ratings for each domain are based on the level of functioning relevant to the circumstances that the subject is in, but the interviewer is required to question in detail to determine how the circumstances arose and whether they were a consequence of the person's own performance.

Most current concepts of personality functioning require that the assessment applies to periods that are free of acute psychiatric disturbance. This expectation is based on the need to differentiate habitual levels of functioning from those that are impaired as a consequence of, say, an episode of depression or an obsessional state. However, this distinction is not always straightforward in practice and it is by no means free of theoretical objections (Docherty *et al*, 1986). Thus, some forms of personality disorder are characterised by recurrent episodes of affective disturbance (this would apply, for example, to so-called borderline personality disorders – Tarnopolsky & Berelowitz, 1987); and others are thought to arise out of, or to constitute part of, a chronic psychiatric condition (as, for

https://doi.org/10.1192/bjp.155.1.24 Published online by Cambridge University Press

instance, with schizophrenia). The empirical testing of associations between personality disorders and other psychiatric conditions is possible only if assessments of both are undertaken separately and if personality functioning is rated both during and apart from episodes of psychiatric disorder. That is the procedure followed by APFA. The interview is designed to be preceded by a lifetime assessment of psychiatric disorder so that periods of functioning that have been accompanied by significant symptomatology can be identified. In our studies we have used a modified version of the SADS–L (Harrington *et al*, 1988).

Interview style

APFA is an investigator-based standardised interview (Brown & Rutter 1966; Graham & Rutter, 1968; Le Couteur *et al*, 1989; Rutter & Brown, 1966). This means that ratings are made by the *interviewer* on the basis of descriptions of behaviour. This contrasts with respondent-based interviews in which ratings are essentially based on whether the *informant* says 'yes' or 'no' to specified questions that ask about particular types of behaviour. The latter work well when there is general agreement on the concept involved but are unsatisfactory when lay and professional concepts differ or where the concept is complex and in need of detailed specification (Breslau, 1987). The advantage of an investigator-based approach in these circumstances is that interviewers can be trained to ensure that the concepts are fully understood and that the codings can provide detailed instructions on what should be taken into account. In this way, it is possible to provide a high degree of structure and operationalisation, but with the structure residing in specification to the interviewer of the concept and coding rather than in specification to the informant of the wording of individual questions. However, for this to be possible it is essential that the concepts be made explicit in the interview schedule and accompanying manual, as they are for APFA, and that interviewers be adequately trained in both the techniques of interviewing and the rules and concepts governing the ratings.

The interviewer's task is to obtain detailed descriptions of the subject's role and activities in specified situations. The onus is on interviewers to use their knowledge of the concept and of the respondent's particular situation to adapt the questions to provide the necessary search for the required information. The ratings in all cases are based on descriptions of behaviour and *not* on affirmative or negative answers to closed questions. Moreover, it is behaviour that is relevant and not attitudes, cognitions or self-concepts. Thus, it is irrelevant whether the subject views him or herself as an effective negotiator or as a good worker or as shy or socially incompetent. Such statements provide the cue to ask for examples of how that is shown, but it is the behaviour and not the attitude that is rated.

Method

Two studies on the qualities of APFA were undertaken:

(a) an inter-rater reliability study, and
(b) an investigation of patterns of malfunction in the different social domains as they related to the hypothesised psychological construct of personality disorder.

Sample

In order satisfactorily to investigate the properties of APFA, it was necessary to have a sample that provided a good spread on quality of social functioning with an adequate number of individuals with pervasive social dysfunction. It was also desirable that the sample covered a wide range of social circumstances in order to test whether the instrument could provide comparable measures when people's living conditions varied greatly. To meet these needs, a sample was constructed from four separate sources. First, there were 40 parents (21 men and 19 women) of children currently attending a psychiatric clinic (the Maudsley Hospital). This is known to be a group with a relatively high rate of psychiatric disorder, often involving chronic social dysfunction (Rutter, 1966, 1987*b*). Second, there were 18 individuals (8 men and 10 women) *currently* attending the adult out-patient department of the same hospital for some non-psychotic disorder involving social difficulties. Third, there were 24 individuals (14 men and 10 women) who were part of a follow-up into adult life of people previously treated for psychiatric problems in childhood. Again, this is a group for which previous research has shown a high rate of psychopathology in adult life (Robins, 1966, 1978; Rutter, 1984). Fourth, there were 15 parents (6 men and 9 women) of these *ex*-patients. In all, this constituted a sample of 97 subjects, of which 49 were male and 48 female.

For some purposes, it was desirable to separate those who were being treated for psychiatric problems in their own right, or who had been so treated in the past. This subgroup was termed 'patients' and was made up of the second and third sources; a total of 42 (22 men and 20 women), with a median age of 31 years (range 24–66). The second subgroup, termed 'parents', comprised the parents of child patients and the parents of ex-child patients, i.e. the first and fourth sources. There were 55 'parents' in all (27 men and 28 women) with a median age of 43 years (range 19–76). Of the total sample, 66 were currently married or cohabiting. Of the remaining 31, 17 had been married previously or had cohabited for at least 6 months. Of the

https://doi.org/10.1192/bjp.155.1.24 Published online by Cambridge University Press

97, 25 had a current psychiatric condition other than personality disorder (assessed on the basis of a SADS–L interview – see Harrington *et al*, 1988) and a further 48 had had such a condition at some time in the past. For obvious reasons, all the subjects who had never had any psychiatric disorder were in the 'parent' group.

Reliability study

For the reliability study, 21 audiotaped interviews were each rated independently by HF, RH and JH (i.e. each interview was rated three times). These 21 interviews comprised 7 interviews by each of the three investigators, which were randomly selected from the 58 interviews from the first two sources (current adult out-patients and parents of current child patients). The sample included 12 men and 9 women, with ages ranging from 26 to 66 years.

The APFA is designed to assess social dysfunction both in relation to specific domains such as work or marriage, and in relation to the hypothesised construct of general dysfunction usually conceptualised as personality disorder. Accordingly, intraclass correlations (Bartko & Carpenter, 1976) were calculated separately for each domain and for the total score (i.e. the sum of the scores for each of the six domains). A particular issue regarding reliability concerns the possible tendency for some interviewers consistently to rate higher or lower than others. This was examined by comparing the mean APFA scores for the three raters, using a one-way analysis of variance.

For each of the social domains, a categorised rating of the type of dysfunction was made (no malfunction, predominant discord/breakdown, predominant avoidance/apathy, and other). For these categorical ratings, inter-rater agreement was assessed using the kappa statistic (Cohen, 1968).

Analysis of patterns of dysfunction

The analysis of patterns of social dysfunction was undertaken on the total sample of 97, but also with separate analyses on groups disaggregated according to patient/parent status and to sex. As already noted, APFA was devised to assess both dysfunction in specific social domains and also generalised social dysfunction. It was assumed that both would occur. Thus, for example, in some cases, individuals with major marital problems will have difficulties that are specific to intense, intimate cohabiting relationships, whereas in other cases the marital problems will form part of a more pervasive pattern of social dysfunction. Accordingly, it was necessary to determine the extent to which the measures for each of the six domains reflected generalised dysfunction – operationally defined as a total APFA score of 16 or greater. This cut-off was chosen because, in order to reach such a score, a person must have had major problems in at least two domains and is likely to have major problems in at least three or four. For this purpose, 'major problems' was defined as a domain subscore of at least 3 (meaning major problems but some significant positive functioning). The value of each domain subscore as a reflection of generalised social dysfunction was assessed in two separate ways. First, the sensitivity and specificity of each subscore were assessed. In this connection, sensitivity constitutes the rate of abnormal subscores in individuals *with* generalised dysfunction; and specificity is the rate of non-abnormal subscores in individuals *without* generalised dysfunction. Second, the confirmatory factor analysis approach of LISREL (Joreskog & Sorbum, 1981) was used to fit a model of a single latent variable of generalised social dysfunction to the data.

Results

Inter-rater reliability

The intraclass correlation coefficients between the three raters for scores on the six separate social domains and for the total APFA score are shown in Table I. Four out of the six domains showed overall correlations that exceeded 0.75, the exceptions being negotiations and everyday coping with correlations of 0.69 and 0.60 respectively. Over the scales as a whole, the raters disagreed by as much as 2 points (on a 6-point scale) in only 6% of cases. It seemed that the somewhat lower, although still acceptable, reliability for the negotiations and everyday coping scales was a function of the rather greater judgement required in rating social functioning in these areas. These two scales also showed the lowest proportion of abnormal scores (28% and 24% respectively versus 30% for work, 57% for love relationships, 47% for friendships, and 37% for non-intimate social contacts in the total sample of 97). The intraclass correlation for the total APFA score was 0.87, with the correlations for individual pairs of raters ranging from 0.85 to 0.90. Agreement between raters on the categorical measure of generalised social dysfunction, operationally defined as a total APFA score of 16 or greater, was also high. There were only three disagreements out of a total of 63 paired comparisons, yielding a kappa value of 0.88.

TABLE I
Agreement between raters 1, 2, and 3 on the score in each social domain and on the total AFPA score (21 subjects)

Social domain	*Rater pairs (ICC[1])*			*All 3 raters*	
	1,2	*1,3*	*2,3*	*ICC*	*Lower 95% CL[2]*
Work	0.74	0.77	0.77	0.76	0.61
Love relationships	0.83	0.79	0.84	0.81	0.68
Friends	0.81	0.76	0.93	0.84	0.73
Non-intimate social contacts	0.89	0.88	0.91	0.89	0.81
Negotiations	(0.64)	0.78	(0.65)	0.69	0.51
Everyday coping	(0.54)	(0.65)	(0.63)	0.60	0.40
Total score	0.85	0.88	0.90	0.87	0.78

1. Intraclass correlation coefficient.
2. Lower 95% confidence limit (for all 3 raters).
All intraclass correlation coefficients were significant at $P < 0.001$ except where bracketed, when $P < 0.01$.

https://doi.org/10.1192/bjp.155.1.24 Published online by Cambridge University Press

The three raters were also generally similar in their level of rating. Thus, the mean total APFA scores for the three raters were 12.4, 12.2 and 12.6, the differences between them falling short of statistical significance on a one-way analysis of variance.

For each social domain with a level rating of 2 or worse, a rating was made of the type of dysfunction (whether discordant or avoidant). The number of ratings was too low to assess reliability separately for each domain (because too many ratings of level of dysfunction were '0' or '1'); accordingly kappa values were calculated pooling the ratings for the six domains. The inter-rater reliability kappas for the three pairs of raters were 0.74, 0.67 and 0.69.

Distribution of scores

There are two different ways in which APFA scores may be used to derive a measure of general social dysfunction. First, the scores of the six separate social domains may be summed to produce a total score with a range from 0 to 30. Second, a count may be made of the number of domains showing major dysfunction (operationally defined as a score of three or greater). Figure 1 shows the distribution of scores for the total sample of 97 according to the first method, and Fig. 2 according to the second method. In this sample of subjects, deliberately weighted to include a relatively high proportion with long-standing psychiatric problems, the modal total score was 10–12, meaning that the average score for each domain showed adequate functioning, but with some problems. Over half the sample (54/97) had scores of 12 or less. Conversely, a third (32/97) had scores of 16 or greater, meaning that there must be major problems in at least two social domains. The picture obtained from looking at the number of domains with major problems (Fig. 2) is broadly comparable. Nearly half (46/97) the sample showed major

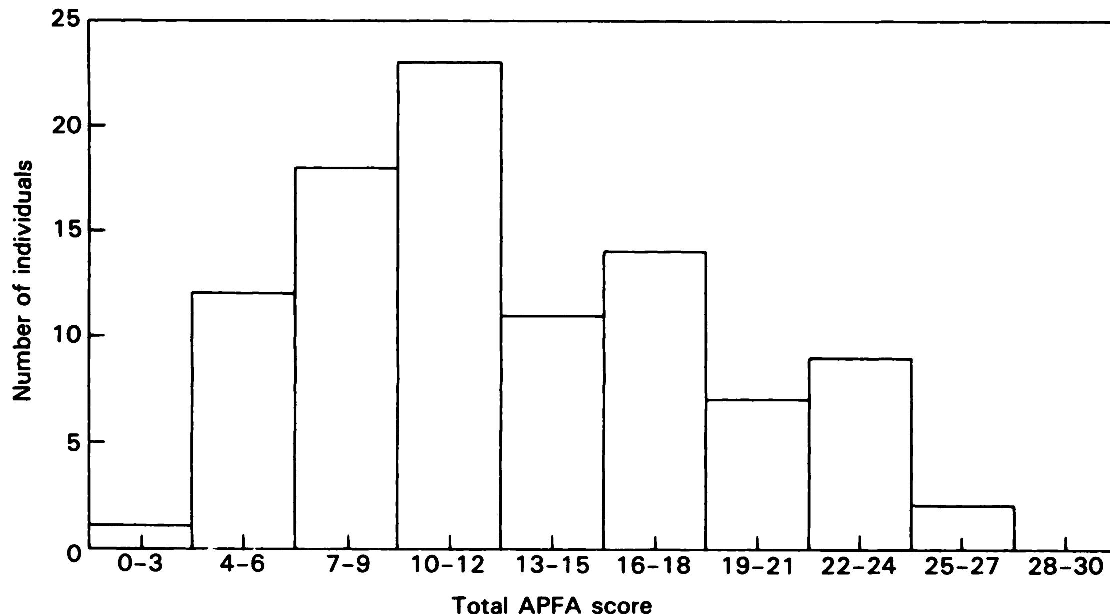

Fig. 1 Distribution of total APFA scores in the whole sample ($n = 97$)

TABLE II

Patterns of social dysfunction: mean (s.d.) domain score according to patient/parent status and according to male/female status

	Work	*Love relationships*	*Friendships*	*Non-intimate social contacts*	*Negotiations*	*Everyday coping*	*Total score*
Patients ($n = 42$)	2.3 (1.4)	3.3 (1.4)	2.6 (1.3)	2.5 (1.3)	2.0 (1.0)	2.2 (1.2)	15.0 (5.9)
Parents ($n = 55$)	1.4 (1.0)	2.5 (1.3)	2.1 (1.4)	1.9 (1.2)	1.8 (1.2)	1.5 (1.0)	11.2 (5.3)
Men ($n = 49$)	1.9 (1.4)	2.8 (1.5)	2.5 (1.4)	2.5 (1.4)	2.1 (1.2)	2.0 (1.1)	13.9 (6.2)
Women ($n = 48$)	1.7 (1.1)	2.8 (1.3)	2.1 (1.3)	1.9 (1.2)	1.7 (1.0)	1.5 (1.1)	11.8 (5.3)

https://doi.org/10.1192/bjp.155.1.24 Published online by Cambridge University Press

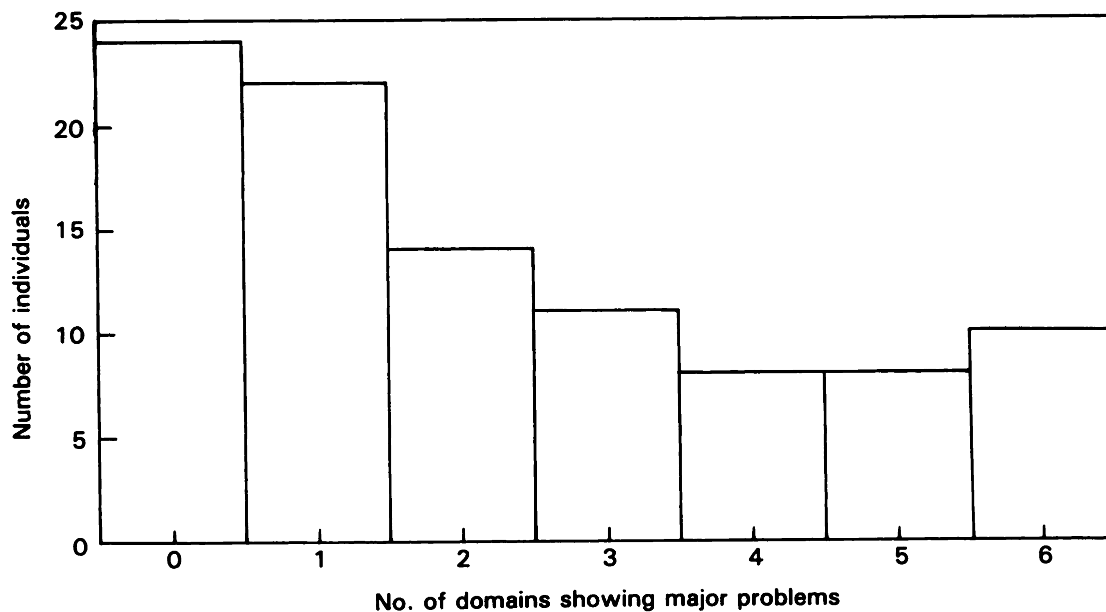

FIG. 2 Distribution of spread of dysfunction across social domains ($n=97$)

problems in only one domain or less, whereas over a quarter (26/97) showed major problems in at least four domains. Not surprisingly, the two approaches tended to pick out the same individuals. All the subjects with major problems in at least four domains had a total APFA score of 16 or greater and of the 32 subjects with a total score of 16 or more, 26 showed major dysfunction in at least four domains.

Table II shows the total scores and separate social domain subscores for patients/ex-patients compared with parents, and for men compared with women. As expected, the mean total score for the patient group (15.1) was substantially higher than that for the parent group (11.2), 45% of the former having scores of 16 or greater compared with 24% of the latter. However, the overall pattern was generally similar in the two groups, with dysfunction most frequent in the domains of marriage and friendships and least frequent in work, negotiations, and everday coping. Similarly, men tended to have higher total scores than women (mean total of 13.7 and 11.9 with 43% and 23% respectively having total of scores of 16 or greater). The only appreciable sex difference in pattern was that women were marginally more likely to show dysfunction in love relationships, whereas they were less likely to show dysfunction in all other domains.

Measurement of general social dysfunction

In order to determine the extent to which dysfunction in each social domain (defined as a subscore of 3 or greater) reflected general social dysfunction (defined as a total score of 16 or greater), sensitivity and specificity coefficients were calculated for each social domain (Table III). There were substantial differences between domains in the pattern shown. Thus, non-intimate social contacts had both high sensitivity (0.87) and specificity (0.88); love relationships had high sensitivity (0.94) but only moderate specificity (0.62); whereas everyday coping had moderate sensitivity (0.67) but very high specificity (0.95).

TABLE III
Sensitivities and specificities of specific social domain dysfunctions as indicators of general social dysfunction ($n=97$)

	Whole sample	*Parents*	*Patients*	*Men*	*Women*
Work					
Sensitivity	0.75	0.54	0.89	0.71	0.82
Specificity	0.92	0.95	0.87	0.93	0.92
Love relationships					
Sensitivity	0.94	0.92	0.95	0.90	1.00
Specificity	0.62	0.64	0.57	0.71	0.54
Friends					
Sensitivity	0.88	0.92	0.84	0.81	1.00
Specificity	0.74	0.76	0.70	0.71	0.76
Non-intimate social contacts					
Sensitivity	0.87	1.00	0.79	0.95	0.73
Specificity	0.88	0.90	0.83	0.82	0.92
Negotiations					
Sensitivity	0.63	0.77	0.53	0.62	0.64
Specificity	0.89	0.86	0.96	0.89	0.89
Everyday coping					
Sensitivity	0.63	0.46	0.74	0.57	0.73
Specificity	0.95	0.95	0.96	0.89	1.00

https://doi.org/10.1192/bjp.155.1.24 Published online by Cambridge University Press

TABLE IV
Associations between dysfunction in specific social domains[1] and general social dysfunction (n = 97)

		Whole sample		Men		Women	
		Absent	Present	Absent	Present	Absent	Present
(a) Dysfunction in love relationships							
Total APFA score	<16	40	25	20	8	20	17
	16+	2	30	2	19	0	11
Sensitivity			0.94		0.90		1.00
Specificity			0.62		0.71		0.54
(b) Dysfunction in everyday coping							
Total APFA score	<16	62	3	25	3	37	0
	16+	12	20	9	12	3	8
Sensitivity			0.63		0.57		0.73
Specificity			0.95		0.89		1.00
(c) Dysfunction in non-intimate social contacts							
Total APFA score	<16	57	8	23	5	34	3
	16+	4	28	1	20	3	8
Sensitivity			0.87		0.95		0.73
Specificity			0.88		0.82		0.92

1. Dysfunction in specific social domain defined as score of 3 or more.

The meaning of these differences is easily appreciated through examination of the cross-tabulations, as shown in Table IV. The high specificity and sensitivity shown by non-intimate social contacts means that nearly all subjects with total APFA scores of 16 or greater showed dysfunction in this domain (28/32), whereas few of those with APFA scores of 15 or less did so (8/65). In other words, the score for this domain was a good indicator of general social dysfunction. Subjects with high APFA scores nearly always showed major problems in love relationships (30/32), but love relationship problems also occurred quite commonly in the absence of general social dysfunction (25/65). Conversely, difficulties in everyday coping rarely appeared as an isolated problem (3/65 subjects had APFA scores of 15 or less). When there was dysfunction in this domain, it was highly likely that there would be general social dysfunction (20/23 subjects) but over a third (12/32) of subjects with APFA scores of 16 or greater did not show dysfunction in everyday coping.

On the whole, these patterns applied similarly to patients and parents and to men and women but there were a few differences in emphasis. Thus, love relationship problems were particularly likely to appear as an isolated dysfunction in women; they did so in three-fifths of subjects (61%) compared with only 30% of men.

LISREL model of personality disorder

The data were also examined to estimate a measurement model for general social dysfunction. The model assumed that the scores from each social domain constituted continuous measures of a single latent variable, namely a hypothetical general social dysfunction factor, but that each score would be subject to 'error' in its reflection of this latent variable. In this connection, 'error' might arise either because the scores also reflected some other latent variable, such as qualities of functioning that were specific to that social domain, or because of measurement bias or unreliability. To achieve a satisfactory fit some correlation among these 'errors' had to be allowed. The eventual chosen model (goodness of fit $\chi^2 = 6.91$ with 6 d.f.) for the whole

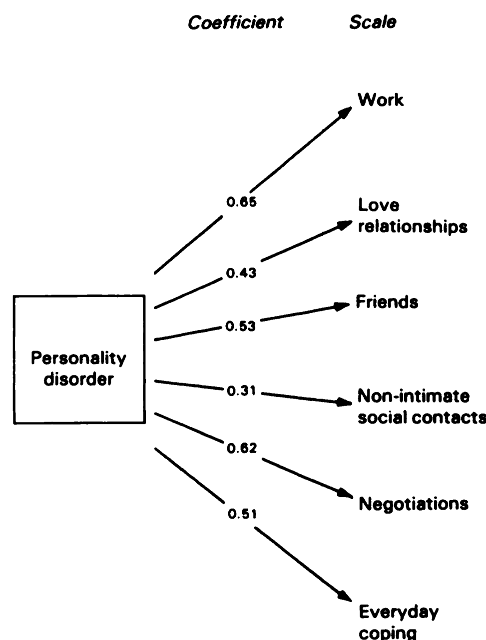

FIG. 3 Estimated measurement model for personality disorder (whole samples $n = 97$).

https://doi.org/10.1192/bjp.155.1.24 Published online by Cambridge University Press

TABLE V
LISREL latent variable coefficients for each social domain

Social domain	*Whole sample*	*Parents*	*Patients*	*Men*	*Women*
Work	0.65	0.51	0.66	0.64	0.68
Love relationships	0.43	0.35	0.46	0.57	0.32
Friends	0.53	0.55	0.55	0.45	0.55
Non-intimate social	0.62	0.65	0.55	0.65	0.50
Negotiations	0.30	0.39	0.24	0.45	0.24
Everyday coping	0.51	0.28	0.60	0.41	0.70

sample is shown in Fig. 3. The coefficients, which measure the degree to which the scale assesses only the latent general social dysfunction variable, confirm the picture obtained from study of sensitivities and specificities. Thus, the coefficients are highest for scores, such as work and non-intimate social contacts, with both high sensitivity and specificity (see Table V). Furthermore, the coefficients for groups disaggregated by sex and by patient/parent status confirmed that the quality of love relationships provided a less good reflection of general social dysfunction in women than in men (0.32 v. 0.57), and that coping assessed general social dysfunction less well in the parents of child-patients than in adults who had themselves been treated for psychiatric disorder (0.28 v. 0.61).

Similarly, although all the scales were intercorrelated through the latent general social dysfunction variable (which accounted for 55% of the variance), there were additional correlations between certain of them, as indicated in Table VI by the estimated correlations between the 'errors'. These correlated 'errors' could represent either artefacts of the rating process or associations across areas in actual functioning. The finding that the residual correlations were markedly different in each of the subgroups (see Table VI) suggests that measurement artefacts are unlikely to constitute the main explanation.

TABLE VI
LISREL analysis: residual correlations between social domains required for a latent variable model of personality disorder

Social domains	*Whole sample*	*Residual correlations*			
		Parents	*Patients*	*Men*	*Women*
Work					
Love relationships	−0.32	−0.48	−0.18	−0.52	−0.08
Friends					
Non-intimate social contacts	0.25	0.05	0.40	0.10	0.13
Negotiations	0.26	0.30	0.26	0.25	−0.05
Everyday coping					

Rather it is probable that the correlations between 'errors' reflect true associations in pattern of social functioning.

Discussion

Our findings clearly indicate that people's social role performance in a range of social domains can be reliably rated using APFA. The intraclass correlation between raters of 0.87 for the total APFA score showed very high reliability and that for individual subscales was also good, with correlations in the 0.80s for three of the six domains and with only one correlation below 0.69 and none below 0.60. The spread of scores was also most satisfactory, with some individuals showing major problems in only one social domain and others doing so in all or almost all domains. The instrument demonstrably works well as a means of measuring the extent of social dysfunction and of identifying different patterns of dysfunction.

The results also show that it is appropriate to regard the total APFA score as a reflection of the degree of general social dysfunction. All the subscales were found to assess the hypothesised construct, or latent variable, of general dysfunction, although they differed in the extent and manner in which they did so. Some, such as 'everyday coping' were highly specific in this connection, so that it was rare for there to be general dysfunction without that domain being affected, but only moderately sensitive so that general dysfunction could be present with only limited dysfunction in that domain. Others, such as 'love relationships', showed the reverse pattern, meaning that specific dysfunction in that domain quite often occurred as an isolated problem but, on the other hand, when there was general social dysfunction it almost always involved impaired love relationships. In principle, it would be possible to create a scoring system for general dysfunction that took account of these differences between domains. However, weightings are notoriously unstable as a result of variations in pattern between different samples. That this would be likely to occur is evident in the finding that, to a minor extent, sensitivities and specificities varied according to patient/parent status and according to sex. Our findings suggest that further consideration of these and other differences (such as age) is warranted. Further empirical work will be necessary to determine how generalisable is the pattern found here. Nevertheless, it would be premature on the basis of this one, relatively small, sample to estimate, and then recommend, the use of particular subjects to group-specific weights or other equivalent changes to the rating procedure.

https://doi.org/10.1192/bjp.155.1.24 Published online by Cambridge University Press

The LISREL modelling brought out another feature; namely that the single latent variable model fitted the data only if a negative residual correlation between the love relationship and friendship scales, and positive residual correlations between non-intimate social contacts and negotiations, and between negotiations and coping, were included. There could be several alternative explanations for these correlations. For example, the need for a negative residual correlation between love relationships and friendship could mean that when there are serious lasting difficulties in love relationships (usually marital discord in the groups studied) occurring as an isolated problem, the individuals tend to seek compensation in friendships. Conversely, those people who have a rewarding committed marital relationship may be less likely to seek extensive extrafamilial friendships. A greater negative correlation was needed in the parents' subgroup than with patients/ex-patients. This could mean that this pattern is most strongly established in older people (as parents tended to be older than patients) or that it is most applicable in individuals who do not suffer from psychiatric disorder. Again, further research is needed to explore these patterns and to determine the mechanisms involved. What is apparent from both the specificity and sensitivity data, as well as from the LISREL modelling, is that there are both specific and general social dysfunctions, and that further study of the associations between the two is needed.

In the meantime, it is reasonable to treat the total APFA score as a fair measure of general social dysfunction. The potential methodological problems stemming from complex, and varying, patterns of associations between subscales are less applicable when total APFA scores are high and when, therefore, by definition there is dysfunction in a majority of the social domains. It is for this reason, as well as others, that there is value in treating high scores, for example, scores of 16 or greater as used in our analyses,* as a categorical indicator of general social dysfunction. It is important, however, to appreciate that this is a matter of practical utility and not of a qualitative distinction. There are numerous hazards in inferring categories or dimensions from distributions (Grayson, 1987). There are many variables that can be regarded as *both* categorical and dimensional, the choice between approaches being based on the purposes for which the measure is being used. Thus, IQ demonstrably functions as a dimensional variable but, equally, severe mental handicap (defined in terms of an IQ below 50) is qualitatively distinct from normal IQ with respect to features such as life expectancy, fecundity and the role of single major gene effects (Rutter & Gould, 1985). But, within the below IQ 50 range, IQ still functions as a dimension for many purposes. The same may well apply to general social dysfunction. Regardless of whether such dysfunction involves qualitative distinctions from normality, it is likely that it may still be present in varying degrees. The present state of affairs is that whereas one may reasonably infer general social dysfunction from high total APFA scores, scores in the intermediate or lower range may reflect lesser degrees of the same variable or, alternatively, higher levels of specific dysfunction. Some guidance is provided by the pattern of APFA scores; thus the former is more likely when there is an even pattern of moderate dysfunction across domains and the latter when the same score is made up of a high score on one or two scales combined with low scores on others. However, the already noted need to take account of the varying pattern of association between subscales means that it is not yet possible to provide a mathematical expression for this differentiation.

The validity of the construct of general social dysfunction may be tested in several different ways. Previous research that has used a concept of personality disorder based on pervasive social dysfunction (but not using APFA as a measure) has shown it to have substantial temporal stability and predictive power (Rutter, 1977; Rutter & Quinton, 1984*a*, 1984*b*; Zeitlin, 1986). The findings suggest that the construct may have validity but it is necessary now to test it more directly using APFA. Such investigations are currently underway with the tests including temporal stability, correlations with childhood behaviour as measured previously in longitudinal studies, correlations with indicators of family psychopathology as studied at an earlier age, family genetic associations, and patterns of concordance in monozygotic and dizygotic twin pairs.

There is also the further question of whether general social dysfunction should be regarded as synonymous with the concept of personality disorder. For obvious reasons, psychiatrists are wary of diagnosing a psychiatric condition merely on the basis of social malfunction. Lewis (1953) clearly pointed out the dangers in doing so. If social malfunction is used as the basis for psychiatric

*As shown by our findings, a cut-off of 16 or greater works reasonably well as a way of selecting individuals with generalised social malfunction. Nevertheless, it is not claimed that this will ultimately prove to be the optimal cut-off point. Further work is necessary to test alternative cut-offs.

https://doi.org/10.1192/bjp.155.1.24 Published online by Cambridge University Press

diagnosis, it is not then possible to argue that a psychiatric condition is the cause of the social failure – that would be a meaningless tautology. It is for these, and similar, reasons that there have been numerous attempts to define personality disorder on the basis of specific psychological dysfunctions such as impulsivity, affective lability, self-destructive behaviour, lack of empathy, sensation-seeking, and lack of remorse (Siever & Klar, 1986). It may well be that the diagnosis of personality disorder should require the presence of such psychological dysfunction as well as impaired social role performance, but it is evident that the use of particular patterns of psychological dysfunction to define supposedly specific types of personality disorder does not work particularly well in practice (Rutter, 1987*a*; Siever & Klar, 1986). The problem is that mixed patterns are the rule rather than the exception. Moreover, both longitudinal and retrospective data suggest that there may be greater coherence in the construct of general social dysfunction than in particular varieties of personality disorder. Thus, conduct disturbance in childhood leads to both antisocial and non-antisocial, somatising and non-somatising varieties of personality disorder in adult life (Quinton *et al*, 1989; Robins, 1986; Zoccolillo *et al*, 1989; Zeitlin, 1986). Similarly, an institutional upbringing does the same (Quinton & Rutter, 1988; Rutter *et al*, 1989). It appears that what are needed are systematic studies to examine empirical associations between psychological traits or symptoms and general social dysfunction. For this purpose, we have undertaken a comparison between APFA scores and findings from a modified version of SADS-L (Harrington *et al*, 1988) as well as the Tyrer Personality Assessment Schedule (Tyrer *et al*, 1979; 1984); the results of that study will be reported separately.

Acknowledgements

This research was supported by a grant from the MacArthur Foundation. Dr Hill was supported by the Medical Research Council on a Training Fellowship. Thanks are due to the Clinicians at the Maudsley Hospital for their collaboration.

References

American Psychiatric Association (1980) *Diagnostic and Statistical Manual of Mental Disorders* (3rd edn) (DSM-III). Washington DC: APA.

Barkto, J. J. & Carpenter, W. T. (1976) On the methods and theory of reliability. *Journal of Nervous and Mental Disease*, **163**, 307-317.

Breslau, N. (1987) Inquiring about the bizarre: false positives in diagnostic interview schedule for children (DISC): ascertainment of obsessions, compulsions and psychotic symptoms. *Journal of the American Academy of Child Psychology*, **26**, 639-644.

Brown, G. W. & Rutter, M. L. (1966) The measurement of family activities and relationships: a methodological study. *Human Relations*, **19**, 241-263.

Clarkin, J., Widiger, T., Frances, A. J., *et al* (1983) Prototypic typology and the borderline personality disorder. *Journal of Abnormal Psychology*, **92**, 263-275

Cohen, J. (1968) Weighted kappa: nominal scale agreement with provision for scaled disagreement of partial credit. *Psychological Bulletin*, **70**, 213-230.

Docherty, J. P., Fiester, S. J. & Shea, T. (1986) Syndrome diagnosis and personality disorder. In *American Psychiatric Association Annual Review*, Vol. 5 (eds A. J. Frances & R. E. Hales). Washington DC: American Psychiatric Association.

Frances, A. J. & Widiger, T. (1986) The classification of personality disorders: an overview of problems and solutions. In *American Pyschiatric Association Annual Review*, Vol. 5 (eds A. J. Frances & R. E. Hales). Washington DC: American Psychiatric Association.

Graham, P. & Rutter, M. (1968) The reliability and validity of the psychiatric assessment of the child. II. Interview with the parent. *British Journal of Psychiatry*, **114**, 581-592.

Grayson, D. A. (1987) Can categorical and dimensional views of psychiatric illness be distinguished? *British Journal of Psychiatry*, **151**, 355-361.

Harrington, R., Hill, J., Rutter, M., *et al* (1988) The assessment of lifetime psychopathology: a comparison of two interviewing styles. *Psychological Medicine*, **18**, 487-493.

John, K. & Weissman, M. M. (1987) The measurement of psychosocial and familial aspects of depression. In *The Measurement of Depression: Clinical, Biological, Psychological and Psychosocial Perspectives* (eds A. J. Marsella, R. M. A. Hirschfeld & M. Ratz), pp. 344-375. New York: Guilford Press.

Joreskog, K. G. & Sorbum, D. (1981) LISREL V: Analysis of linear structural relationships by maximum likelihood and least squares methods. *Research Report 81-88*. University of Uppsala.

Katschnig, H. (1983) Methods for measuring social adjustment. In *Methodology in Evaluation of Psychiatric Treatment* (ed. T. Helgason), pp. 205-218. New York: Cambridge University Press.

Le Couteur, M., Rutter, M., Lord, C., *et al* (1989) Autism diagnostic interview: a standardized investigator-based instrument. *Journal of Autism and Developmental Disorders* (in press).

Lewis, A. J. (1953) Health as a social concept. *British Journal of Sociology*, **4**, 109-124.

Mann, A. H., Jenkins, R., Cutting, J. C., *et al* (1981) The development and use of a standardized assessment of abnormal personality. *Psychological Medicine*, **11**, 839-847.

Mellsop, G., Varghese, F., Joshua, S., *et al* (1982) The reliability of axis 2 of DSM-III. *American Journal of Psychiatry*, **139**, 1360-1361.

Merikangas, K. & Weissman, M. M. (1986) Epidemiology of DSM-III Axis II: personality disorders. In *American Psychiatric Association Annual Review*, Vol. 5. (eds A. J. Frances & R. E. Hales), pp. 258-278. Washington DC: American Psychiatric Association.

Quinton, D. & Rutter, M. (1988) *Parenting Breakdown: The Making or Breaking of Intergenerational Links*. Aldershot: Gower Publishing.

——, —— & Gulliver, L. (1989) The social functioning in early adulthood of the children of psychiatric patients. In *Straight and Devious Pathways From Childhood to Adulthood* (eds L. N. Robins & M. Rutter). New York: Cambridge University Press (in press).

Reich, J. (1985) Measurement of DSM-III, Axis II. *Comprehensive Psychiatry*, **26**, 352-363.

Robins, L. N. (1966) *Deviant Children Grown Up*. Baltimore: Williams and Wilkins.

—— (1978) Sturdy childhood predictors of adult antisocial behaviour: replications from longitudinal studies. *Psychological Medicine*, **8**, 611-622.

https://doi.org/10.1192/bjp.155.1.24 Published online by Cambridge University Press

—— (1986) The consequences of conduct disorder in girls. In *Development of Antisocial and Prosocial Behavior: Research, Theories and Issues* (eds D. Olweus, J. Block & M. Radke-Yarrow), pp. 385–414. New York: Academic Press.

RUTTER, M. (1966) *Children of Sick Parents: An Environmental and Psychiatric Study*. Institute of Psychiatry Maudsley Monograph No 16. London: Oxford University Press.

—— (1977) Prospective studies to investigate behavioural change. In *The Origins and Course of Psychopathology* (eds J. S. Strauss, H. M. Babigian & M. Roff), pp. 223–247. New York: Plenum Publishing.

—— (1984) Psychopathology and development. I. Childhood antecedents of adult psychiatric disorder. *Australian and New Zealand Journal of Psychiatry*, **18**, 225–234.

—— (1987*a*) Temperament personality and personality disorder. *British Journal of Psychiatry*, **150**, 443–458.

—— (1987*b*) Parental mental disorder as a psychiatric risk factor. In *American Psychiatric Association Annual Review*, Vol. 6 (eds R. E. Hales & A. J. Frances). Washington DC: American Psychiatric Association.

—— & BROWN, G. W. (1966) The reliability and validity of measures of family life and relationships in families containing a psychiatric patient. *Social Psychiatry*, **1**, 38–53.

—— & QUINTON, D. (1984*a*) Parental psychiatric disorder: Effect on children. *Psychological Medicine*, **14**, 853–880.

—— & —— (1984*b*) Long-term follow-up of women institutionalized in childhood: factors promoting good functioning in adult life. *British Journal of Developmental Psychology*, **2**, 191–204.

—— & GOULD, M. (1985) Classification. In *Child and Adolescent Psychiatry: Modern Approaches* (2nd edn) (eds M. Rutter & L. Hersov) pp. 304–321. Oxford: Blackwell Scientific.

—— & PICKLES, A. (1989) Improving the quality of psychiatric data: classification, cause and course. In *Methodological Issues in Longitudinal Research I: Data and General Designs* (ed D. Magnusson). Cambridge: Cambridge University Press (in press).

——, QUINTON, D. & HILL, J. (1989) The effects of an institutional rearing on social functioning in early adult life in men and women. In *Straight and Devious Pathways from Childhood to Adulthood* (eds L. N. Robins and M. Rutter). Cambridge: Cambridge University Press (in press).

SIEVER, L. J. & KLAR, H. (1986) A review of DSM–III criteria for the personality disorders. In *American Psychiatric Association Annual Review*, Vol. 5 (eds A. J. Frances & R. E. Hales), pp. 279–314. Washington DC: American Psychiatric Association.

STANGL, D., PFOHL, B., ZIMMERMAN, M., *et al* (1985) A structured interview for the DSM–III personality disorders: a preliminary report. *Archives of General Psychiatry*, **42**, 591–596.

STURT, E. & WYKES, T. (1987) Assessment schedules for chronic psychiatric patients. *Psychological Medicine*, **17**, 485–494.

TARNOPOLSKY, A. & BERELOWITZ, M. (1987) Borderline personality: a review of recent research. *British Journal of Psychiatry*, **151**, 724–734.

TYRER, P. & FERGUSON, B. (1987) Problems in the classification of personality disorder. *Psychological Medicine*, **17**, 15–20.

——, ALEXANDER, M. S., CICCHETTI, D., *et al* (1979) Reliability of a schedule for rating personality disorder. *British Journal of Psychiatry*, **135**, 168–174.

——, CICCHETTI, D. V., CASEY, R. R., *et al* (1984) Cross-national reliability study of a schedule for assessing personality disorders. *Journal of Nervous and Mental Disease*, **172**, 718–721.

WEISSMAN, M. M. (1975) The assessment of social adjustment: a review of techniques. *Archives of General Psychiatry*, **32**, 357–365.

——, SHOLOMSKAS, D. & JOHN, K. (1981) The assessment of social adjustment: an update. *Archives of General Psychiatry*, **38**, 1250–1258.

WORLD HEALTH ORGANIZATION (1978) *International Classification of Diseases* (9th revision). Geneva: WHO.

ZEITLIN, H. (1986) *The Natural History of Disorder in Childhood*. Institute of Psychiatry Maudsley Monograph No 29. London: Oxford University Press.

ZOCCOLILLO, M., RUTTER, M., QUINTON, D., *et al* (1989) The outcome of conduct disorder: implications for defining adult personality disorders (in preparation).

*J. Hill, MRCP, MRCPsych, *Senior Lecturer*; R. Harrington, MRCPsych, MPhil, *Lecturer, Institute of Psychiatry*; H. Fudge, BA, *Research Worker, Institute of Psychiatry*; M. Rutter, CBE, MD, FRCPsych, FRS, *Honorary Director, MRC Child Psychiatry Unit*; A. Pickles, PhD, *Statistician, MRC Child Psychiatry Unit*

*Correspondence: *Department of Child and Adolescent Psychiatry, Royal Liverpool Children's Hospital, Alder Hey, Eaton Road, Liverpool L12 2AP*

https://doi.org/10.1192/bjp.155.1.24 Published online by Cambridge University Press