Using the SWAP-200 in a personality-disordered forensic population: is it valid, reliable and useful?

LUISA E. MARIN-AVELLAN[1] GILL MCGAULEY[1, 2], COLIN CAMPBELL[1] AND PETER FONAGY[3, 4], [1]Department of Mental Health, St George's Hospital Medical School, London; [2]Department of Forensic Psychotherapy, West London Mental Health NHS Trust, Broadmoor Hospital, Crowthorne, Berkshire, UK; [3] Sub-Department of Clinical Health Psychology, University College London, UK; [4]Anna Freud Centre, London, UK

ABSTRACT

Background *Treatment and risk management of forensic patients relies heavily on diagnosing psychopathology, yet the reliability of clinical diagnoses of personality disorder has been found to be only fair to low. Structured instruments for the global assessment of personality disorder are infrequently used in clinical assessments possibly due to their limited validity and clinical utility.*

Aims/methods *The Shedler-Westen Assessment Procedure-200 (SWAP-200) was developed in an effort to address these limitations. Although good reliability and validity in relation to clinicians' diagnosis of personality disorder has been reported, to date the validity of this instrument has not been assessed in relation to other standardized instruments or in a personality-disordered, forensic population. This study aims to establish the reliability and validity of the SWAP-200 against other diagnostic instruments and measures of interpersonal functioning in a personality disordered forensic population.*

Results *This paper reports the results of 30 subjects from a high secure hospital in the UK who were assessed with the SWAP-200, the Structured Clinical Interview for DSM-IV Personality Disorders (SCID-II), the Adult Attachment Interview (AAI) and the Chart of Interpersonal Reactions in Closed Living Environments (CIRCLE). Preliminary results suggest that the SWAP-200 is a reliable instrument for the diagnosis of personality disorder in forensic patients.*

Conclusions *Although the small sample size allows only preliminary conclusions about the validity of this instrument, early results show a reduction of the diagnosis of comorbidity compared with the SCID-II, together with an increased number of expected associations between independent measures of interpersonal functioning and categories of personality disorder.*

Introduction

Clinical work with forensic patients relies heavily on diagnosing psychopathology to aid risk management and treatment planning, yet the reliability of clinical diagnoses of personality disorder has shown to be only low to fair (Mellsop et al., 1982; Shea et al., 1990). Although standardized diagnostic instruments that globally assess personality disorder, such as the Structured Clinical Interview for DSM-IV, Axis II Personality Disorders (SCID-II) (First et al., 1997); the International Personality Disorder Examination (IPDE) (Loranger et al., 1994); the Structured Interview for DSM-IV Personality Disorders (SIDP-IV) (Pfohl et al., 1994); and the Diagnostic Interview for DSM-IV (DIPD-IV) (Zanarini et al., 1994) have shown improved diagnostic reliability (Zimmerman, 1994), a survey of forensic services in the United Kingdom revealed that the use of these instruments in clinical assessments is rare (Milton, 2000).

In relation to validity, Blackburn (Blackburn, 2000) notes the low levels of agreement reported between measures (Hyler, Skodol et al., 1990; O'Boyle and Self, 1990; Skodol et al., 1991; Pilkonis et al., 1995). Additionally, poor agreement has been found between patients and informants regarding the presence and types of personality disorder, informants reporting significantly more pathology than patients (Zimmerman et al., 1988; Dowson, 1992; Mann et al., 1999). It has been suggested that the main factor limiting the validity of current diagnostic instruments is that they rely heavily on information about the personality obtained from questionnaires or interviews whose structure leaves little room for clinical judgement (Blackburn, 2000; Tyrer, 2000). This method of assessment may be flawed when used with personality-disordered patients due to limitations in their capacity for introspection and perspective taking, considered by some clinicians as a diagnostic criterion for personality disorder (Westen and Shedler, 1999a). Furthermore, the tendency of direct questioning to elicit socially desirable answers from patients regarding their personality has been cited as a factor affecting the validity of the information (Zimmerman, 1994), especially in forensic settings (Hiscoke et al., 2003).

Concerns regarding the discriminant validity of these instruments arise as a result of the high rates of diagnostic comorbidity found between Axis I and Axis II disorders of the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) (American Psychiatric Association, 1994), but especially within Axis II disorders (Blackburn, 2000). In offender populations in England, including patients detained in secure facilities under the Legal Category of Psychopathic Disorder (Mental Health Act, 1983) and violent male prisoners, 95% have been found to meet criteria for more than one personality disorder; the average number of diagnoses reported was 3.6 disorders per subject (Coid, 1992). Reported disorders spanned all Axis II categories (Coid, 1992; Blackburn and Coid, 1999; Blackburn et al., 2003).

High rates of multiple diagnoses could be explained by possible flaws in the classification system of personality disorder. Research findings indicate that individual criteria do not clearly discriminate between DSM categories of personality disorder (Blais and Norman, 1997) and factor analysis fails to confirm that these diagnostic criteria are organized into the categories proposed (Hyler, Lyons et al., 1990). In addition, Westen and Shedler (1999a) have commented that the DSM artificially dichotomizes continuous variables into present and absent and that the narrow set of diagnostic criteria used does not encompass all domains of personality functioning, thus failing to capture the complexity of clinical data.

The Shedler–Westen Assessment Procedure-200 (SWAP-200) (Westen and Shedler, 1999a, 1999b) was developed in an effort to address the limitations of current diagnostic instruments and to provide clinically meaningful descriptions of personality disorder. Although good reliability and validity in relation to clinicians' diagnosis of personality disorder has been reported (Westen and Shedler, 1999a), to date the validity of this instrument has not been assessed either in relation to other standardized instruments or in forensic populations.

This article reports the preliminary results of a reliability and validity study of the SWAP-200 compared with the SCID-II, a widely used diagnostic instrument for the assessment of personality disorder (Blackburn, 2000), in a population of personality-disordered forensic patients. However, owing to the centrality of interpersonal functioning within the conceptualization of personality disorder, we also included in the analysis two observer measures related to interpersonal functioning for further validation of the SWAP-200, the Adult Attachment Interview (AAI) (George et al., 1984, 1996) and the Chart of Interpersonal Reactions in Closed Living Environments (CIRCLE) (Blackburn and Renwick, 1996). The AAI assesses an individual's current state of mind with regard to attachment (Main and Goldwyn, 1994/1998). Although not a direct measure of interpersonal functioning, attachment representations affect expectations of and drive behaviour towards others (Dozier and Tyrrell, 1998).

The study addressed the following questions:

(1) Can satisfactory inter-rater reliability for the SWAP-200 be achieved in a forensic, personality-disordered population? How does inter-rater reliability for the SWAP-200 compare with that of the SCID-II?
(2) Is the discriminant validity of the diagnosis of personality disorder improved using the SWAP-200 compared with the SCID-II, as indicated by a reduction in the diagnosis of comorbidity between Axis II disorders?
(3) As a further assessment of validity, do the SWAP-200 categories of personality disorder better describe dimensions of interpersonal functioning, as measured by the AAI and the CIRCLE, compared with the SCID II?

Method

Sample

The participants were 30 admissions to a high secure hospital in the UK. The selection criteria of subjects were being 18 years and over and the presence of personality disorder according to DSM-IV research criteria. Patients whose active symptoms of psychosis or whose proficiency in English precluded them completing the instruments or who had a physical disability or an IQ below 70 were excluded.

The mean age of the sample was 30.3 years (SD = 8.04; range: 19–50), 73% were male and the mean IQ was 91.38 (SD = 15.76; range: 70–135). In terms of legal classification under the Mental Health Act (1983), 33% fell under the category of Mental Illness, while 53% were classified under Psychopathic Disorder; the rest had a dual categorization. The SCID-I diagnoses (First et al., 1997) are summarized in Table 1.

Regarding criminological history, 80% of the sample had previous convictions and had served up to 16.4 years in prison (mean: 4.3 years; SD = 4.9). Table 2 summarizes the past offending history and index offences of the subjects.

Table 1: SCID-I diagnoses	
SCID-I categories	Percentages
Mood Disorders	37
Schizophrenia	30
Other Psychoses	23
Alcohol Abuse	53
Alcohol Dependence	33
Drug Abuse	53
Drug Dependence	30
Anxiety Disorders	13

Table 2: Offending history		
Categories of offence	Previous offences (%)	Index offence (%)
Major violence[1]	25	54
Minor violence[2]	66	17
Sexual offences	17	13
Arson	33	25
Criminal damage	42	8
Acquisitive offences	71	21
Other	43	10

Notes: [1] Includes homicide, attempted murder, infanticide and grievous bodily harm. [2] Includes actual bodily harm, assault, making an affray, wounding and threats of violence.

Procedure

After securing Local Research Ethical Approval, potential subjects were identified from sequential admissions to the hospital. Written informed consent was obtained and the participants were assessed during the first four months of their admission with the instruments listed below.

Instruments

(1) The SCID-II (First et al., 1997) is a semi-structured interview that assesses the presence or absence of the 10 DSM-IV personality disorders (Avoidant, Dependent, Obsessive-Compulsive, Paranoid, Schizotypal, Schizoid, Histrionic, Narcissistic, Borderline, Antisocial Personality Disorder), and the appendix categories, Depressive Personality Disorder, Passive-Aggressive Personality Disorder and Personality Disorder Not Otherwise Specified. Items are coded as: Insufficient Information, Absent or False, Subthreshold and Threshold or True. Diagnosis is reached by the presence of at least the minimum predetermined number of criteria for each of the categories. Following Blackburn and Coid (1999), dimensionalized scores, the sum of all criteria within each category, were also derived.

(2) The SWAP-200 (Westen and Shedler, 1999a) is a Q-sort instrument designed to quantify clinical judgement of personality pathology. The set of 200 personality-descriptive statements is ranked into eight categories by a clinician with a thorough knowledge of the patient, according to how well they describe a particular patient but following a fixed distribution. The resulting ordering of the items is then compared with diagnostic prototypes representing each DSM Axis II personality disorders to ascertain the degree of match. The correlation coefficient obtained for each category is transformed into T scores with a mean of 50 and a standard deviation of 10. The cut-off point for the diagnosis of personality disorder is T = 60 (Shedler, personal communication). The diagnostic prototypes were obtained by asking a large sample of clinicians in the United States of America to describe prototypical, hypothetical patients representing each category of personality disorder in its 'pure' form. The resulting SWAP descriptions were averaged to arrive at a single, aggregate prototype representing the core clinical consensus on the features of each personality disorder (Westen and Shedler, 1999a). Overall, these diagnostic prototypes were found to be different from DSM criteria.

(3) The AAI (George et al., 1984 and 1996) is a semi-structured in-depth interview designed to assess the mental representation of attachment in adults. It asks a series of open-ended questions about early attachment relationships and experiences with significant attachment figures, about separations, rejection, loss, trauma and physical and sexual abuse. Interviews are audio-

recorded and transcribed verbatim and are then rated by a detailed discourse analysis. There are three reliable classifications of attachment in the Main and Goldwin system (Main and Goldwyn, 1991). Individuals are described as Secure-Free-Autonomous (F) with regard to attachment if they value attachment relationships and are able to describe them objectively, irrespective of whether these experiences were negative or positive; Insecure-Dismissive (Ds) if they are dismissing, devaluing or cut off from attachment relationships and experiences; and Insecure-Preoccupied (E) if they are confused, unobjective and preoccupied with or by past attachment relationships and experiences.

(4) The CIRCLE (Blackburn and Renwick, 1996) is a 49-item set that assesses social behaviour of psychiatric inpatients. Each item is assigned to one of the eight scales measuring the octants of the interpersonal circle denoted as Dominant, Coercive, Hostile, Withdrawn, Submissive, Compliant, Friendly and Sociable. Items are rated on a four-point scale (from 0 to 3) by an observer who has known the patient for at least one month. The CIRCLE was derived empirically from the interpersonal circle (Wiggins, 1982), which is an empirical framework based on Leary's (Leary, 1957) interpersonal theory for describing interpersonal variables.

Rating the instruments

SCID II was administered and coded by two pairs of raters, two psychiatrists and one psychologist. Two of these raters remained the same for the coding of the SWAP-200, together with a new rater who was a psychiatrist. The clinical experience since qualification of these three raters ranged from 4.5 to 16 years. A new procedure for rating the SWAP-200 was developed using the audio-recording and typed transcript of the AAI along with a standard set of information comprising the admission psychiatric, psychological and social work reports.

The AAI was administered by four clinicians and researchers trained in the use of the instrument. The transcribed AAI's were rated for the two-way, Secure versus Insecure, and three-way, Secure (F), Insecure-Dismissing (Ds) and Insecure-Preoccupied (E) classification (Main and Goldwyn, 1994). The two AAI coders were trained by Mary Main and Erik Hesse and confirmed as reliable for the two- and three-way classification.

The CIRCLE was completed by two nurses who had known the subject for at least one month and could rate their interpersonal functioning on the ward during the previous week. The two ratings were done sequentially over a two-week period; the final score was the average of the two ratings. Each nurse was blind to the other's rating and had no access to the rest of the information collected for the research. In addition, the nurses were not involved in writing the clinical reports that were used for the rating of the SWAP-200.

Results

Inter-rater reliability

Inter-rater reliability was assessed for 10 cases. The average kappa values for the SCID-II between two pairs of raters was 0.84 with a range from 0.44 to 1. In relation to the SWAP-200, the average Pearson coefficient for the three raters was 0.89, with a range from 0.81 to 0.96.

Diagnoses using the SCID-II and the SWAP-200:

Frequency of diagnostic overlap within Axis II

The SCID-II results (Figure 1) showed that all the categories of personality disorder were represented, with the exception of Histrionic which was absent and Schizotypal which was only present in one case. The most prevalent diagnoses were Avoidant, Depressive, Paranoid, Borderline and Antisocial Personality Disorder, and to a lesser extent Passive-Aggressive and Schizoid Personality Disorder. The average number of diagnostic categories met by the sample was 3.3 disorders per subject (range 1–7).

The SWAP-200 results (Figure 2) showed that the most prevalent diagnoses were Paranoid, Borderline and Antisocial, followed by Schizotypal and Narcissistic Personality Disorder. The categories of Avoidant, Dependent, Obsessive-Compulsive, Depressive and Histrionic were either absent or only present in one case. The average number of diagnostic categories met by the sample was 2.4 disorders per subject (range 0–6).
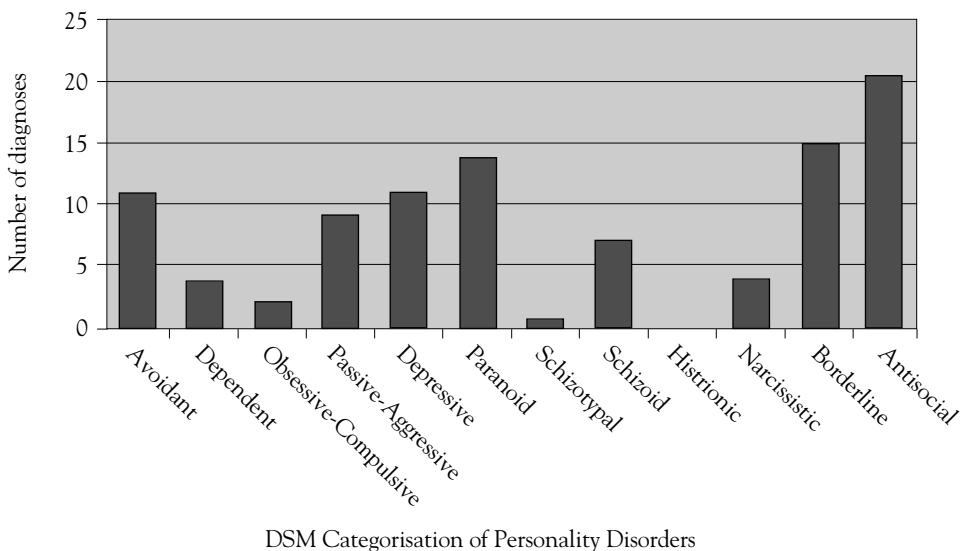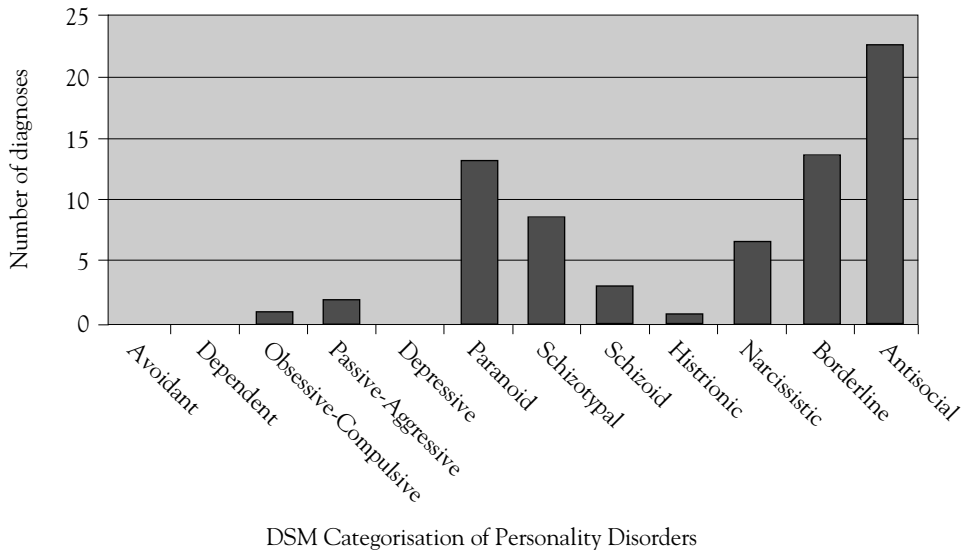

Figure 1: SCID-II diagnoses.

Figure 2: SWAP-200 diagnoses.

We computed the associations between SCID-II and SWAP-200 diagnoses using both a categorical approach, indicated by the presence or absence of a disorder, and a dimensional approach comparing the sum of scores in each category of the SCID-II with T-scores obtained from the SWAP-200 (Table 3). The level of agreement between equivalent categorical diagnoses on SCID-II and SWAP-200 was generally low due to the fact that the SWAP-200 decreased the frequency of diagnoses compared with the SCID-II and that it diagnosed fewer categories of personality disorder. From the remaining categories represented in the SWAP-200 results, Kappa values reached significance only for two categories, Narcissistic and Antisocial Personality Disorder. However, the dimensional results revealed a higher level of agreement for Antisocial Personality Disorder and showed a moderate but significant level of agreement for all the categories with the exception of Schizotypal and Narcissistic.

Associations between attachment classifications and the diagnosis of personality disorder

Categorical diagnosis of personality disorder using the SCID-II and the SWAP-200

The relationship between categorical diagnoses of personality disorders and attachment classification was marginal, as the SCID-II revealed only a significant association between Schizoid personality disorder and an Insecure-Dismissing classification of attachment and no significant associations were found between SWAP-200 categorical diagnoses and attachment classifications.

Table 3: Agreement of SWAP-200 diagnoses with regard to SCID-II

Categories of personality disorder	Agreement of categorical diagnoses (Kappa)	Agreement of dimensional measurements of PD (Pearson)
Avoidant	cc	0.44*
Dependent	cc	0.41*
Obsessive-Compulsive	0.65	0.49**
Passive-Aggressive	0.10	0.39*
Depressive	cc	0.46*
Paranoid	0.13	0.48**
Schizotypal	−0.06	−0.07
Schizoid	0.30	0.52**
Histrionic	cc	0.43*
Narcissistic	0.54**	0.33
Borderline	0.27	0.37*
Antisocial	0.49**	0.73**

Notes: cc = cannot be calculated; * = $p<0.05$; ** = $p<0.01$.

Dimensional measurement of personality disorder

Table 4 displays the point-biserial correlations computed between selected SCID-II and SWAP-200 categories and the AAI classifications. The SCID-II showed only one significant association in the two-way analysis of attachment classifications (Secure versus Insecure) for Avoidant Personality Disorder and none in the three-way analysis.

Table 4: Associations between AAI classifications & dimensional measurements on SCID-II and SWAP-200

Categories of personality disorder		AAI classifications			
		Secure (n = 4)	Insecure (n = 26)	E (n = 7)	Ds (n = 19)
SCID-II	Avoidant		0.43**		
SWAP–200	Avoidant		0.34*		
	Obsessive-Compulsive				0.34*
	Paranoid			0.33*	
	Schizotypal		0.35*		
	Schizoid		0.38*		0.38*

Notes: * = $p<0.05$; ** = $p<0.01$; E = Insecure-Preoccupied; Ds = Insecure-Dismissing.

SWAP-200 results yielded significant positive associations between scores on the Avoidant, Schizotypal and Schizoid Personality Disorder dimensions and Insecure attachment in the two-way analysis of attachment classifications. In the three-way analysis, the Insecure-Preoccupied classification was associated with higher scores on the Paranoid Personality Disorder dimension and the Insecure-Dismissing classification was associated with higher scores on the dimension of Obsessive-Compulsive and Schizoid Personality Disorder. Thus, results show a greater number of associations between the SWAP-200 proto-types and attachment classifications than between SCID-II categories and attachment classifications.

Associations between styles of interpersonal functioning and the diagnosis of personality disorder

Categorical diagnosis of personality disorder using the SCID-II and the SWAP-200

We also computed the association between the CIRCLE measures of overt styles of interpersonal functioning and the assessment of personality disorder using the listed approaches. Using categorical diagnoses on the SCID-II, six of the DSM categories were positively associated with CIRCLE dimensions. However, the level of associations was very limited and did not discriminate well between disorders. Avoidant was linked to Submissive, Depressive was associated with Hostile and Submissive, Schizoid was associated with non-Dominant,

Table 5: Associations between the CIRCLE and categorical diagnoses of PD on SCID-II and SWAP-200

Categories of personality disorders		Dominant	Coercive	Hostile	Withdrawn	Submissive	Compliant	Friendly	Sociable
SCID-II	Avoidant					0.40*			
	Depressive			0.39*		0.34*			
	Schizoid	−0.37*							
	Narcissistic	0.36*							
	Borderline		0.44**						
	Antisocial	0.42*							
SWAP-200	Schizoid	−0.34*							
	Narcissistic	0.37*			−0.45**	−0.40*			
	Borderline	0.39*	0.39*						0.36*

*Notes: * = p<0.05; ** = p<0.01*

Borderline was liked to Coercive and Narcissistic and Antisocial personality disorders were both characterized as Dominant.

The SWAP-200 showed fewer but a more differentiated set of associations with the CIRCLE styles, as Narcissistic Personality Disorder is described as Dominant but also Non-Withdrawn and Non-Submissive, and Borderline is described as Dominant, Coercive and Sociable. Similarly to the SCID-II, Schizoid Personality Disorder was associated only with low scores in the Dominant style on the CIRCLE.

Dimensional measurements of personality disorder

The use of the dimensional scores in the SCID-II shows a better differentiated set of associations compared with the categorical diagnoses, as the dimension of Depressive is described as Hostile and Submissive, Schizoid is Withdrawn and Submissive, and Histrionic is described as Coercive, Hostile and Non-Compliant. In addition, the dimensions of Avoidant, Passive-Aggressive and Narcissistic Personality Disorder were positively associated with CIRCLE dimensions, although only to a very limited extent.

Table 6: Associations between the CIRCLE and dimensional measurements of PD on SCID-II and SWAP-200

Categories of personality disorders		CIRCLE							
		Dominant	Coercive	Hostile	Withdrawn	Submissive	Compliant	Friendly	Sociable
SCID-II	Avoidant					0.54**			
	Passive-Aggressive			0.41*					
	Depressive			0.38*		0.48*			
	Schizoid				0.58**	0.68**			
	Histrionic		0.50**	0.53**			−0.39*		
	Narcissistic	0.46*							
SWAP-200	Avoidant	−0.46*				0.58**	0.48*		
	Dependent	−0.42*				0.53**	0.39*		
	Obsessive-Compulsive	−0.44*							
	Depressive	−0.43*				0.57**			
	Schizoid	−0.42*	−0.39*			0.43*	0.39*		
	Histrionic	0.48*							
	Narcissistic	0.52**			−0.47*	−0.55**			
	Borderline	0.39*	0.45*						
	Antisocial	0.59**	0.39*		−0.40*	−0.57**	−0.39*		

Notes: * = p<0.05; ** = p<0.01.

The SWAP-200 shows a more differentiated set of associations with the CIRCLE styles, as nine dimensions of personality disorder had correspondence with different styles of interpersonal functioning measured by the CIRCLE. The dimensions Avoidant, Dependent and Schizoid personality disorder are described as Non-Dominant, Submissive and Compliant, although Schizoid is the only one that was differentiated from the other two dimensions by negative scores in the Coercive style on the CIRCLE. Depressive is characterized as Non-Dominant and Submissive. The Narcissistic and Antisocial prototypes are both described as Dominant, Non-Withdrawn and Non-Submissive, although Antisocial is differentiated by being Coercive and Non-Compliant, which corresponds to the descriptors of this disorder. In addition, the Borderline dimension of personality disorder is described as Dominant and Coercive; Obsessive-Compulsive is associated with Non-Dominant and Histrionic with Dominant.

Discussion

The study sample is similar to those reported in previous studies of high secure forensic patients in terms of the prevalence of Axis I and Axis II disorders (Coid, 1992; Blackburn and Coid, 1999; Blackburn et al., 2003). Our results using the SWAP-200 agree with previous findings in relation to the prominence of Antisocial, Borderline, Paranoid and Narcissistic Personality Disorder. However, they also show a lower frequency of Schizoid and a higher prevalence of Schizotypal Personality Disorder compared to previous studies (Coid, 1992; Blackburn et al., 2003).

The main aims of the study were to determine the reliability and some types of validity of the SWAP-200 in forensic populations. The latter is a diagnostic instrument for personality disorder based on clinical knowledge about the personality. It uses a statistical procedure to improve the reliability of clinical judgements.

Our first hypothesis was that the SWAP-200 would show satisfactory inter-rater reliability in a forensic, personality-disordered population compared with the SCID-II. The results confirmed this hypothesis (average Kappa values for the SWAP-200 were 0.89, as opposed to 0.84 for the SCID-II) and replicated previously reported high values by the authors of the SWAP-200 (Westen and Shedler, 1999a). Of note was the finding that there was little variability in the range of inter-rater reliability Kappa values on the SWAP-200, compared with the SCID-II (0.81 to 0.96 on the SWAP-200, as opposed to 0.44 to 1 on the SCID-II). This was achieved despite the fact that there were (a) differences in the raters' professional backgrounds and years of clinical experience and that (b) the rating of the SWAP-200 is based entirely on clinical judgement as opposed to patients' self-reported symptoms during a semi-structured interview. Research has shown that diagnosis by clinical judgement has only low to moderate reliability (Mellsop et al., 1982; Shea et al., 1990). Our findings of a uniformly high

inter-rater reliability suggest that the problem might lie not in the use of clinical observations and inferences, but in the procedure that is used to reach diagnosis (i.e. unstructured vs. statistical).

The second hypothesis was that the SWAP-200 would reduce the frequency with which subjects were diagnosed with multiple categories of personality disorder compared with the SCID-II, because diagnosis using the SWAP-200 is based on a wider selection of criteria and on a distribution of personality traits, as opposed to exclusive criteria for each disorder. Our results confirmed this hypothesis as the average rate of diagnostic comorbidity was reduced from 3.3 to 2.4 personality disorders per subject. In addition, the SWAP-200 generated a reduced span across the range of personality disorder, as results were mainly distributed across seven categories, i.e. Passive-Aggressive, Paranoid, Schizotypal, Schizoid, Narcissistic, Borderline and Antisocial, compared with the diagnostic spread of the SCID-II results which spanned all the categories with the exception of Schizotypal and Histrionic Personality Disorder. It is therefore unsurprising that the level of agreement between these two diagnostic instruments was low for most of the categories, with the exception of Antisocial and Narcissistic Personality Disorder, which showed moderate agreement. However, the dimensional use of the SWAP-200 and the SCID II showed additional moderate agreement for all the categories of personality disorders, with the exception of Schizotypal and Narcissistic.

The above findings suggest that the diagnostic prototypes of personality disorders in the SWAP-200 have some similarities with DSM descriptors measured by the SCID-II but that the pathway to diagnosis is reached differently with the two instruments. In addition, it could be argued that when used categorically the SCID-II might over-rate personality traits associated with these disorders, when in fact they are only present as dysfunctional traits that do not reach the threshold of a personality disorder. An important characteristic of the diagnostic procedure of the SWAP-200 is that it is not the presence or absence of a particular personality trait that determines the diagnosis of a category, but the constellation of personality traits resulting from the ranking of the 200 traits. Consequently two or more diagnostic categories can share similarly high ranked traits, such as a restricted range of expression of emotions (only characteristic of Schizoid Personality Disorder in DSM-IV categorization) or the tendency to feel victimized and to be suspicious of others' intentions (only characteristic of Paranoid Personality Disorder). However, the meaning of these items will be determined by the context of the other items that the clinician ranked as being most descriptive of the patient. This diagnostic approach aims to minimize the frequency of comorbid diagnoses without compromising the complexity of clinical data. Also, as items from different prototypes of personality pathology can coexist in the clinician's rating of the SWAP-200, elevations in specific dimensions of personality disorder can occur without necessarily reaching the diagnostic cut-off point for a personality disorder.

   As our results showed lower rates of diagnostic comorbidity using the SWAP-200 compared with the SCID-II it is possible that the SWAP-200 might have under-diagnosed personality disorder. Furthermore, a recent study using the SWAP-200 reported disagreements between patients' self-report and clinicians' judgement of personality disorder and concluded that there is no certainty about who is more accurate at describing personality dysfunction (Davidson et al., 2003). In the absence of a gold standard against which to assess the validity of diagnoses based on the SWAP-200, the study included two measures related to interpersonal functioning (the AAI and the CIRCLE). These were rated by independent informants who were blind to the SCID-II and SWAP-200 results. We expected that the results regarding interpersonal functioning would be associated with specific categories of personality disorder in a way that corresponded to the descriptors of these disorders. Our third hypothesis was that the SWAP-200 results would be better associated with the results generated by these independent measures compared with the SCID-II. Whilst the general direction of associations for both instruments was similar, the SWAP-200 yielded more detailed and expected associations between the different AAI categories and the CIRCLE, and several categories of personality disorder compared with the SCID-II. However, the use of dimensional measurements of personality disorder proved to be better in terms of finding a wider range of associations with both the AAI and the CIRCLE, which is in agreement with previous findings regarding the advantages of using dimensional measures as opposed to categorical instruments (Widiger and Frances, 1994).

   The literature reports that insecure attachment classifications are over-represented in forensic patients with personality disorder (van Ijzendoorn et al., 1997). Consequently the finding that, even with a small sample, AAI classifications differentiated five of the DSM prototypes of personality disorder diagnosed by the SWAP-200 compared with only one category diagnosed by the SCID-II is a promising finding with regard to the validity of the SWAP-200. Although research has reported only a few specific associations between insecure attachment classifications and categories of personality disorder (Fonagy et al., 1996; Dozier et al., 1999) this study confirmed two of these despite the over-representation of the Insecure-Dismissing category (Ds), namely between Schizoid Personality Disorder and Insecure-Dismissing and between Paranoid Personality Disorder and Insecure-Preoccupied classification of attachment.

   The CIRCLE styles of interpersonal functioning differentiated categories of personality disorder diagnosed with the SWAP-200 to a greater extent than those diagnosed with the SCID-II. Furthermore, previously reported findings suggesting that personality disorders appear more differentiated along the dimension of Power (i.e. Dominance-Submission) than along the Affiliation axis are supported by the current results (Blackburn, 1998). Although the number of associations with SWAP-200 diagnoses is modest compared with previously reported findings using the MCMI-I (Blackburn, 1998), correlation

coefficients are higher. Also, the fact that significant associations were found in spite of the small sample size (n = 30 compared with n = 104; Blackburn, 1998) is an encouraging finding. The pattern of correlations found was highly congruent for the different dimensions of personality disorder, especially for Schizoid, Narcissistic and Antisocial Personality Disorder diagnosed by the SWAP-200.

The authors acknowledge the following limitations of the study. First, the small sample size does not allow us to draw as robust conclusions about our hypotheses when compared with studies based on larger samples. Second, although an increased number of significant associations were found between SWAP-200 diagnoses and independent measures of interpersonal functioning compared with the SCID-II, these results are still modest. However, in view of the fact that many associations were significant in spite of the small size of the sample, we would expect to find a larger number of significant associations with a larger sample of subjects. Third, this study used a new rating procedure for the SWAP-200 based on audio-taped AAI interviews and review of admission clinical reports, as opposed to the systematic clinical interview for the SWAP-200. Consequently, the reliability of this new rating procedure still needs to be established. It is anticipated that these limitations will be addressed on completion of the full study as this will include an additional 40 subjects (final n = 70) and a reliability study on the new rating procedure for the SWAP-200.

Conclusions

The preliminary results of this study are encouraging in terms of establishing the reliability and validity of the SWAP-200 in a personality-disordered forensic population. Of particular importance are the findings that this instrument decreased the frequency of diagnostic overlap of personality disorder categories, i.e. diagnostic comorbidity, compared with the SCID-II and showed more significant associations with other instruments that measure or relate to interpersonal functioning. The potential implications of these findings are twofold. First, reliability results of the SWAP-200 show that it can be successfully used to structure clinical diagnosis. These results suggest that it is the procedure for reaching a diagnosis (i.e. statistical as opposed to unstructured) rather than the use of data collected through structured-self-report interviews that improves the diagnosis of personality disorder. Second, our preliminary finding of a decreased frequency of diagnostic comorbidity of personality disorder in this population using a clinically anchored diagnostic instrument opens the possibility of further assessing the contribution of personality factors in treatment and risk-management planning.

Acknowledgements

References

American Psychiatric Association (1994) *Diagnostic and Statistical Manual of Mental Disorders*, 4th edn. Washington, DC: Author.

Blackburn R (1998) Relationship of personality disorders to observer ratings of interpersonal style in forensic psychiatric patients. *Journal of Personality Disorders* 12: 77–85.

Blackburn R (2000) Classification and assessment of personality disorders in mentally disordered offenders: a psychological Perspective. *Criminal Behaviour and Mental Health* 10: S8–S32.

Blackburn R, Coid J (1999) Empirical clusters of DSM-III personality disorders in violent offenders. *Journal of Personality Disorders* 13: 18–34.

Blackburn R, Logan C, Donnelly J, Renwick S (2003) Personality disorders, psychopathy and other mental disorders: co-morbidity among patients at English and Scottish high security hospitals. *Journal of Forensic Psychiatry & Psychology* 14: 111–137.

Blackburn R, Renwick S (1996) Rating scales for measuring the interpersonal circle in forensic psychiatric patients. *Psychological Assessment* 8: 76–84.

Blais MA, Norman DK (1997) A psychometric evaluation of the DSM-IV personality disorder criteria. *Journal of Personality Disorders* 11: 168–176.

Coid J (1992) DSM-III diagnosis in criminal psychopaths: a way forward. *Criminal Behaviour and Mental Health* 2: 78–94.

Davidson K, Obonsawin M, Seils M, Patience L (2003) Patient and clinician agreement on personality using the SWAP-200. *Journal of Personality Disorders* 17: 208–218.

Department of Health (1983) *Mental Health Act 1983*. London: Stationery Office.

Dozier M, Stoval KC, Albus KE (1999) Attachment and psychopathology in adulthood. In Cassidy J, Shaver PR, eds. *Handbook of Attachment*. New York: Guilford Press, pp. 497–519.

Dozier M, Tyrrell C (1998) The role of attachment in therapeutic relationships. In Simpson JA, Rholes WS, eds. *Attachment Theory and Close Relationships*. New York: Guilford Press, pp. 221–248

Dowson JH (1992) Assessment of DSM-III-R personality disorders by self-report questionnaires: the role of informants and a screening test for comorbid personality disorders. *British Journal of Psychiatry* 161: 344–352.

First MB, Spitzer RL, Gibbon M, Williams, JBW (1997) *Structured Clinical Interview for DSM-IV Axis I Disorders – Clinician Version (SCID-I)*. Washington DC: American Psychiatric Press.

First MB, Gibbon M, Spitzer, RL, Williams JBW, Benjamin LS (1997) *Structured Clinical Interview for DSM-IV Axis II Disorders (SCID-II)*. Washington DC: American Psychiatric Press.

Fonagy P, Leigh T, Steele M, Steele H, Target M, Gerber A (1996) The relation of attachment status, psychiatric classification, and response to psychotherapy. *Journal of Consulting and Clinical Psychology* 64: 22–31.

George C, Kaplan N, Main M (1984) The adult attachment interview. Privileged communication, University of California at Berkeley.

George C, Kaplan N, Main M (1996) The Adult Attachment Interview, 3rd edn. Unpublished manuscript, University of California at Berkeley.

Hiscoke UL, Langstrom N, Ottosson H, Grann M (2003) Self reported personality traits and disorders (DSM-IV) and risk of criminal recidivism: a prospective study. *Journal of Personality*

Disorders 17: 293–305.

Hyler SE, Lyons M, Rieder R, Young L, Williams J, Spitzer R (1990) The factor structure of self-report DSM-III axis II symptoms and their relationship to clinicians' ratings. *American Journal of Psychiatry* 147: 751–757.

Hyler SE, Skodol AE, Kellman D, Oldham JM, Rosnick L (1990) Validity of the Personality Diagnostic Questionnaire-Revised: a comparison with two structured interviews. *American Journal of Psychiatry* 147: 1043–1048.

Leary T (1957) *Interpersonal Diagnosis of Personality*. New York: Ronald Press.

Loranger AW, Sartorius N, Andreoli A, Berger P, Buchheim P, Channabasavanna SM, Coid B, Dahl A, Diekstra RFW, Ferguson B, Jacobsberg LB, Mombour W, Pull C, Ono Y, Reiger DA (1994) The International Personality Disorder Examination: the WHO/ADAMHA International Pilot Study of Personality Disorders *Archives of General Psychiatry* 51: 215–224.

Main M, Goldwyn, R (1991) Adult Attachment Classification System, Version 5. Unpublished manuscript, University of California at Berkeley.

Main M, Goldwyn, R (1994; 1998) Adult Attachment Scoring and Classification System. Unpublished manuscript, University of California at Berkeley.

Mann AH, Raven P, Pilgrim S, Khanna S, Velayudham A, Suresh KP, Channabasavanna SM, Janca A, Sartorius N (1999) An assessment of the Standardized Assessment of Personality as a screening instrument for the International Personality Disorder Examination: a comparison of informant and patient assessment for personality disorder. *Psychological Medicine* 29: 985–989.

Mellsop G, Varghese F, Joshua S, Hicks A (1982) The reliability of Axis II of DSM-III. *American Journal of Psychiatry* 139: 1360–1361.

Milton, J. (2000) A postal survey of assessment procedures for personality disorder in forensic settings. *Psychiatric Bulletin* 24: 254–257.

O'Boyle M, Self D (1990) A comparison of two interviews for DSM-III-R personality disorders. *Psychiatry Research* 32: 85–92.

Pfolh B, Blum N, Zimmerman M, Stangl D (1994) *Structured Interview for DSM-IV Personality (SIDP-IV)*. Iowa City: University of Iowa Department of Psychiatry.

Pilkonis PA, Heape CL, Proietti JM, Clark SW, McDavid JD, Pitts TE (1995) The reliability and validity of two structured diagnostic interviews for personality disorders. *Archives of General Psychiatry* 52: 1025–1033.

Shea MT, Pilkonis PA, Beckham E, Collins JF, Elkin I, Sotsky SM, Docherty JP (1990) Personality disorders and treatment outcome in the NIMH Treatment of Depression Collaborative Research Program. *American Journal of Psychiatry* 147: 711–718.

Skodol AE, Oldham JM, Rosnick L, Kellman HD, Hyler SE (1991) Diagnosis of DSM-III-R personality disorders: issues in assessment. *International Journal of Methods in Psychiatric Research* 1: 13–26.

Tyrer P (2000) Improving the assessment of personality disorders. *Criminal Behaviour and Mental Health* 10: S51–S65.

Van IJzendoorn MH, Feldbrugge JTTM, Derks FCH, de Ruiter C, Verhagen MA, Philipse MWG, van der Staak CPF, Riksen-Walraven JMA (1997) Attachment representations of personality disordered criminal offenders. *American Journal of Orthopsychiatry* 67: 449–459.

Westen D, Shedler J (1999a) Revising and assessing axis II, Part I: Developing a clinically and empirically valid assessment method. *American Journal of Psychiatry* 156: 258–272.

Westen D, Shedler J (1999b) Revising and assessing axis II, Part II: Toward an empirically based and clinically useful classification of personality disorders. *American Journal of Psychiatry* 156: 273–285.

Widiger TA, Frances AJ (1994) Toward a dimensional model for the personality disorders. In: Costa PT, Widiger TA, eds. *Personality Disorders and the Five-Factor Model of Personality*. Washington, DC: American Psychological Association, pp. 19–39.

Wiggins JS (1982) Circumplex models of interpersonal behaviour in clinical psychology. In Kendal PC, Butchers JN, eds. *Handbook of Research Methods in Clinical Psychology*. New York:

Wiley, pp.183–221.

Zanarini MC, Frankenburg FR, Sickel AE, Yong L (1994) *Diagnostic Interview for DSM-IV Personality Disorders (DIPD-IV)*. Belmont, MA: McLean Hospital.

Zimmerman M, Pfohl B, Coryell W, Stangl D, Corenthal C (1988) Diagnosing personality disorders in depressed patients: a comparison of patient and informant interviews. *Archives of General Psychiatry* 45: 733–737.

Zimmerman M (1994) Diagnosing personality disorders. *Archives of General Psychiatry* 51: 225–245.

Address correspondence to: Luisa E. Marin-Avellan, Research Fellow, Department of Mental Health, Jenner Wing, St George's Hospital Medical School, London SW17 0RE, UK. Email: lavellan@sghms.ac.uk