The Achenbach System of Empirically Based Assessment as a measure of personality disorder traits in children and adolescents

Ioanna Vrouva

D.Clin.Psy. Thesis (Volume 1), 2012

University College London

Overview

Many psychological problems in adulthood have their roots in childhood and adolescence. This is particularly true for personality disorders (PDs). In order to identify young people with PD traits before their problems become pervasive, we need reliable and valid assessment tools. This volume includes three papers seeking to examine the usefulness of the Achenbach System of Empirically Based Assessment (ASEBA) for measuring PD traits in young people.

Part 1 is a systematic review of cross-sectional and longitudinal studies that used the ASEBA to investigate the internalising and externalising problems of young people presenting with (or who later developed) personality difficulties. The majority of the studies examined antisocial and borderline PD. The review concluded that there was consistent evidence of criterion validity for a few ASEBA scales but the ASEBA did not have adequate psychometric properties for accurately identifying young people with PD.

Part 2 is an empirical paper that used a large database created for audit purposes in a community-based psychotherapy and counselling service for young people. The ASEBA profiles of young people with PD traits and PD-related presenting problems were examined. This paper also describes the development and psychometric evaluation of two new, PD-related ASEBA scales.

Finally, Part 3 is a critical appraisal of the research undertaken. It discusses epistemological and methodological aspects of the work and reflects upon the proposed changes in the conceptualisation of PD in the updated diagnostic system. This paper also highlights the clinical dilemmas related to diagnosing PD before adulthood.

TABLE OF CONTENTS

Contents of Tables and Figures

Acknowledgements

Firstly I wish to thank my UCL supervisor Professor Peter Fonagy for all his input and support throughout this project, and for all the intellectual stimulation and personal kindness I have received from him over the years. I will miss his humour that could turn even the most disappointing statistical figure into a source of laughter and inspiration.

I would also like to cordially thank my external supervisor Dr Geoffrey Baruch and his staff at the Brandon Centre for supporting this project in every way. Dr Baruch provided his assistance beyond the call of duty on more than one occasion and his consistently prompt responses to my questions were hugely appreciated.

Many thanks are also due to Dr Anthony Bateman and Dr Marco Chiesa for their willingness to provide expert ratings, and to Professor Chris Barker and Professor Nancy Pistrang for research teaching and project support of exceptional quality.

My friends Leslie Muetzelfeldt, Aleksandra Urbanska, and Eleana Stathopoulou helped me format this volume at short notice and deserve special thanks for their generosity.

I am also very indebted to Dr Adam Duncan for his insightful comments on earlier drafts of the manuscript and for all the empowering vibes he unfailingly sent throughout the writing process.

Finally I wish to thank my family, friends, mentors and clients for supporting this work in their own unique way.

Part 1: Literature Review

How useful is the Achenbach System of Empirically Based Assessment for measuring personality disorder traits in children and adolescents?

A systematic literature review

Abstract

Aim: To systematically evaluate the psychometric properties of the Achenbach System of Empirically Based Assessment (ASEBA) as a potential measure of personality disorder (PD) traits in children and adolescents.

Method: 29 studies published in the last decade are systematically reviewed in a narrative synthesis. To be included, studies had to have used at least one scale of the Child Behaviour Checklist (CBCL; Achenbach, 1991), the Teacher's Report Form (TRF; Achenbach, 1991), or the Youth Self-Report (YSR; Achenbach, 1991), and to have investigated at least one PD, measured categorically (as a diagnostic entity) or dimensionally (e.g. PD traits questionnaire).

Results: Most studies focused on Antisocial PD (ASPD)/psychopathy, or Borderline PD (BPD). There was evidence of criterion validity; ASEBA externalising scales were associated with ASPD and psychopathy, and both internalising and externalising scales were associated with BPD. Furthermore, the CBCL-Dysregulation Profile had modest predictive validity. The validity estimates reported were widely varying, depending on methodological issues such as design (cross-sectional or longitudinal), and shared method variance.

Conclusions: The literature to date does not provide a compelling case for use of the ASEBA as a tool with adequate psychometric properties for assessing PD traits in juveniles. For the time being, the ASEBA can be used only tentatively to inform clinicians about PD traits in young people who may require more focused assessment. There is a need for large-scale, preferably prospective research to explore the reliability and validity of specific ASEBA scales and item sets as measures of personality pathology in young people.

Introduction

The notion that some individuals have maladaptive personality characteristics dates back to Hippocrates's body humours. However, the research enterprise on personality pathology was launched more than 2000 years later.

The Diagnostic Statistical Manual (DSM) Fourth Edition Text Revision (DSM-IV-TR) defines personality disorder (PD) as: *'an enduring pattern of inner experience and behaviour that deviates markedly from the expectations of the individual's culture*, *is pervasive and inflexible, has an onset in adolescence or early adulthood, is stable over time, and leads to distress or impairment'* (APA, 2000, p. 685). These patterns may manifest as persistent disturbances in cognition, affect, interpersonal functioning, and impulse control (Crawford et al., 2008). The DSM-IV lists ten PD types, which are grouped into three Clusters. These are listed in Table 1, alongside the pervasive patterns defining them.

Table 1. *The DSM-IV personality disorders and the pervasive patterns that define them* (*as presented by Clark, 2009, p. 28*)

Personality disorder type	Pervasive patterns
Cluster A (*odd or eccentric*)	
Paranoid (PPD)	Distrust and suspiciousness of others such that their motives are interpreted as malevolent
Schizoid (SZPD)	Detachment from social relationships and a restricted range of expression of emotions in interpersonal settings
Schizotypal (STPD)	Social and interpersonal deficits marked by acute discomfort with, and reduced capacity for, close relationships, as well as by cognitive or perceptual distortions and eccentricities of behaviour
Cluster B (*dramatic, emotional or erratic*)	
Antisocial (ASPD)	Disregard for and violation of the rights of others occurring since age 15 years
Borderline (BPD)	Instability of interpersonal relationships, self-image, and affects, and marked impulsivity
Histrionic (HPD)	Excessive emotionality and attention seeking
Narcissistic (NPD)	Grandiosity (in fantasy or behaviour), need for admiration, and lack of empathy
Cluster C (*anxious or fearful*)	
Avoidant (AvPD)	Social inhibition, feelings of inadequacy, and hypersensitivity to negative evaluation
Dependent (DPD)	Excessive need to be taken care of that leads to submissive and clinging behaviour and fears of separation
Obsessive-compulsive (OCPD)	Preoccupation with orderliness, perfectionism, and mental and interpersonal control, at the expense of flexibility, openness, and efficiency

PDs are associated with life-course psychosocial dysfunction in a wide range of domains, including limited, unstable, or maladaptive interpersonal relationships, social isolation, poor occupational performance, ineffective coping strategies, interpersonal violence, and suicide (e.g., NIMHE, 2003; Skodol et al., 2002). Moreover, people with PD

are more vulnerable to other mental health problems such as depression and substance misuse, and have a poorer prognosis for treatment of these disorders (Bender et al., 2001).

Although the historical belief was that PDs should not be diagnosed prior to adulthood, growing evidence suggests that juvenile PD is a valid clinical concept (Crick, Murray-Close, & Woods, 2005; Miller, Muehlenkamp, & Jacobson, 2008; Sharp & Romero, 2007). It has been argued that like adults, some young people's personality difficulties cause sufficient impairment to necessitate psychological treatment (Shiner, 2009). Therefore, reliable and valid assessment tools are needed for research and clinical purposes.

Achenbach's system of empirically based assessment (ASEBA; www.aseba.org) is one of the most extensively used and well-validated sets of instruments for the broadband screening of emotional and behavioural problems. In recognition of the importance of including information from multiple respondents in the assessment of children (Achenbach, McConaughy, & Howell, 1987) several forms exist, including the Child Behaviour Checklist (CBCL; Achenbach, 1991), the Teacher's Report Form (TRF; Achenbach, 1991) and the Youth Self-Report (YSR; Achenbach, 1991).

The items comprising the ASEBA are organised into eight narrowband syndromes, namely: Withdrawn/Depressed, Somatic Complaints, Anxious/Depressed, Social Problems, Thought Problems, Attention Problems, Delinquent Behaviour, and Aggressive Behaviour. There are also two broadband scales: a) Internalising Problems, which include the Withdrawn/Depressed, Somatic Complaints and Anxious/Depressed scales, and b) Externalising problems, which include the Delinquent and Aggressive Behaviour scales. These are aggregated in a Total Problems score, which is an average of all syndrome scales. Achenbach and associates have also created Axis I DSM-oriented scales, which partition the ASEBA items in different ways compared to the statistically-based syndromes.

Although Achenbach and colleagues have not created DSM-oriented scales for PDs, several scales and items in the ASEBA forms conceptually relate to personality trait characteristics and could therefore apply to Axis II diagnoses (Kernberg, Weiner, & Bardenstein, 2000). This is in keeping with evidence suggesting that specific behavioural or emotional symptoms traditionally described within the internalising-externalising spectrum of psychopathology are significant childhood risk factors for later personality dysfunction (Cohen, 2008; De Clercq & De Fruyt, 2007).

However, the extent to which the ASEBA may be used to assess PD traits in young people is an empirical question. Because the ASEBA is among the most widely employed instruments in the assessment of children's mental health problems, it was expected that there would be a sufficient number of studies that have used the ASEBA to investigate juvenile PD. The overarching aim of this review is a comprehensive examination and synthesis of these studies' findings, in order to evaluate the psychometric properties of the ASEBA forms as potential measures of PD traits in children and adolescents. Specifically, this review seeks to answer the following questions:

a) Are ASEBA scales associated with categorical and dimensional measures of PD? Which scales have the strongest concurrent validity?
b) Can ASEBA scales that were completed in childhood or adolescence significantly predict PD in adulthood? Which scales have the strongest predictive validity, and what is the relevant sensitivity/specificity/predictive value?

Method

Search strategy

Studies were identified from searches up to May 2012 in four databases: PsycINFO, MEDLINE, EMBASE, and CINHAL. Electronic searches were based on both subject headings terms and textwords. Truncated and adapted terms were used to allow for variations in American/English spelling. All PD types were the subject of this review,

and the following search terms were used to identify PD: *character disorder** or *character dysfunct** or *character pathology* or *personality disorder** or *personality dysfunct** or *personality pathology* or *PD* or *Axis II* or *BPD* or *ASPD*. Adjacency operators were used (within two words apart) to combine the nine DSM-IV PDs (i.e. *schizoid* or *schizotypal* or *paranoid* or *narcissis** or *histrion** or *borderline* or *antisocial* or *obsessive-compulsive* or *avoidant* or *dependent*) with the terms *personality* or *PD*. Adjacency operators within two words apart were also used to combine the terms *depressive*, *passive-aggressive*, *sadoma?ochistic*, *ma?ochistic*, *self-defeating*, (which reflect earlier conceptualisations of PD) with the terms *personality* or *PD*.

To identify ASEBA scales, the terms *ASEBA, Achenbach*, *CBCL*, *Child Behavi*r Checklist*, *YSR*, *Youth Self-Report*, *TRF*, *Teacher* Report Form*, *C-TRF*, and *Caregiver-Teacher* Report Form* were used.

The next step to the search involved combining terms for PDs with those for the ASEBA scales. The domains searched were title, abstract, heading word, table of contents, key concepts, original title, and tests and measures. At this stage, a few limits were applied to the results. For both cultural relevance and practical considerations, the search was restricted to articles published between 2002 and 2012. Studies undertaken in any country were included, provided that the article was available in English and was published in a peer-reviewed journal.

Inclusion and Exclusion criteria

Only articles providing the data necessary for the calculation of the correspondence between at least one PD and one ASEBA scale were included, resulting in the following inclusion criteria: (a) an empirical study based on original data collection; (b) inclusion of data on at least one ASEBA narrowband or broadband scale (i.e. from the CBCL, TRF, or YSR) and at least one PD; (c) PD measurement including either a standard diagnosis derived from a structured interview or a clinician-based diagnosis, or

questionnaire measures of PD completed by children and young people themselves, clinicians or significant others. Both cross-sectional and longitudinal studies were included.

Although psychopathy is outside the official diagnostic nomenclature (Johnstone & Cooke, 2004), studies that focused on psychopathy were included. This is because it is widely recognised that psychopathy is a PD, characterised by manipulativeness, superficial charm, egocentricity, shallow affect, lack of remorse, unreliability, and impulsivity (Hare, 2003).

The exclusion criteria were: (a) all qualitative studies, single case studies and predominantly theoretical papers; (b) studies which did not include an independent criterion for PD (i.e. other than the ASEBA); (c) studies focusing solely on related or frequently comorbid psychological problems (e.g. self-harm) but not PDs; (d) studies which provided descriptive ASEBA data (e.g. mean scores) for a PD group but without presenting equivalent data for a non-PD group; (e) studies which presented descriptive PD and ASEBA data but without an analysis of the association between the two; (f) studies using other ASEBA scales designed for younger (such as the CBCL/1.5- 5), or older (such as the Young Adult Self-Report) age groups. For parsimony, studies exploring the link between ASEBA and maladaptive personality traits associated with PD, such as neuroticism and impulsivity, were also excluded.

Study selection process

Initially, the search was limited by publication year (2002 onwards), publication type (peer reviewed journals) and language (English). This generated 1562 citations, of these PsycINFO generated 589, MEDLINE generated 240, EMBASE generated 275 and CINAHL generated 458. Following an initial screen for obvious duplicates, the total of number of publications included for the initial screening was 1113.

At this stage the main criteria used was whether the study a) included both a PD and an ASEBA scale and b) whether it used an appropriate study population. As decisions

were made on the basis of titles and abstracts alone, they erred on the side of caution. After screening of titles and abstracts, 455 studies merited closer inspection. After review of full text articles, 417 citations were removed because they did not meet the inclusion criteria, and a further twelve studies could not be retrieved. The reference lists of identified studies were also hand searched. Reference tracking identified three additional studies resulting in a final total of 29 studies for review. The above selection process is summarised in Figure 1.

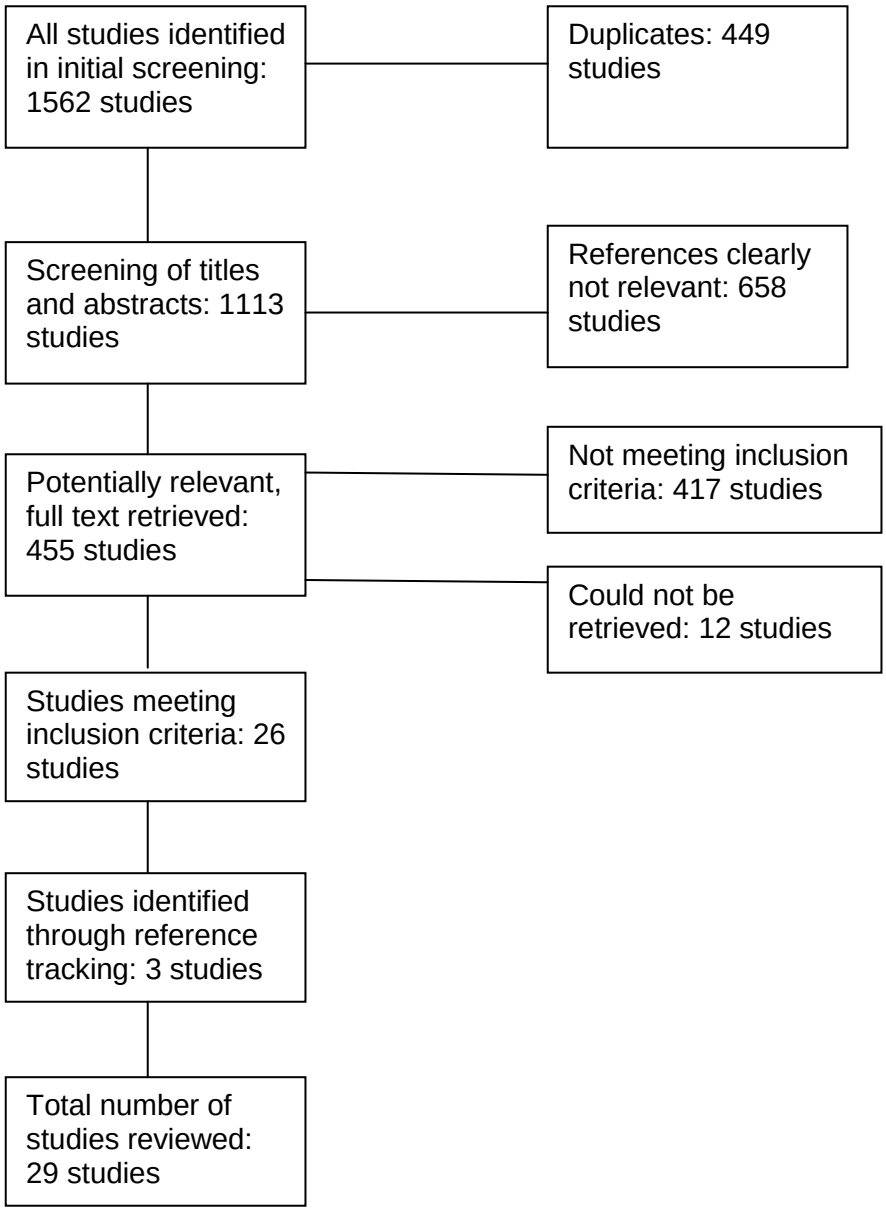

Figure 1. Flowchart of study selection

Results

The 29 citations included in the review originated from 22 research projects and are outlined in Table 2 along with their main characteristics.

Table 2. *Characteristics of included studies*

Study	Cohort	Country	Sample: *N*, age at first (ASEBA) assessment (mean, SD) % female % Caucasian	Sample type	ASEBA form(s) used	ASEBA scales used	PD type studied	PD criterion used	Design
Arens, Grabe, Spitzer, & Barnow (2011)	SHIP (Study of Health in Pomerania)	Germany	*N*=68 Mean (SD) age=15.3 (2.2) 82% female ?% Caucasian	A clinical group with BPD & age, sex & education matched clinical controls (DEP) & healthy controls	YSR, CBCL (mother)	Internalising Externalising	BPD	SCID-II	Longitudinal (5 years follow-up)
Barnow, Lucht, & Freyberger (2005)	SHIP (Study of Health in Pomerania)	Germany	*N*=168 Mean (SD) age=14.5 (2.1) 48% female 100%Caucasian	Healthy controls & a group with positive family history of alcoholism	YSR, CBCL (mother)	Aggr.Behav. Delin. Beh.	ASPD	SCID-II (for partic/ants >15 years old)	Cross-sectional
Burnette & Reppucci (2009)	Virginia female detainees	USA	*N*=121 Mean (SD) age=16.2 (1.3) 100% female 38% Caucasian	Incarcerated girls	YSR	Aggr.Behav.	BPD	SIDP-IV	Cross-sectional
Burnette, South, & Reppucci (2007)	Virginia female detainees	USA	*N*=121 Mean (SD) age=16.2 (1.3) 100% female 38% Caucasian	Incarcerated girls	YSR	Aggr.Behav.	Cluster B	SIDP-IV	Cross-sectional
Carlson, Egeland, & Sroufe (2009)	Minnesota longitudinal study of risk and adaptation	USA	*N*=162 age≅12 years, 49% female 67% Caucasian	Community poverty sample of at risk first-born children of young mothers	TRF (teacher)	28 items representing instability & disturbance in emotional, attentional, behavioural, & relational domains	BPD	SCID-II	Longitudinal (16 years follow-up)

Dolan & Rennie (2006)	North West England detainees	UK	*N*=115 Mean (SD) age= 16.1 (.9), 0% female 86% Caucasian	Incarcerated adolescent males with conduct disorder	CBCL (parent)	All	Psychopathy	PCL:YV	Cross-sectional
Dolan & Rennie (2007)	North West England detainees	UK	*N*=115 Mean (SD) age= 16.2 (.9), 0% female 86% Caucasian	Incarcerated adolescent males with conduct disorder	CBCL (parent)	All	Psychopathy	YPI	Cross-sectional
Dutra, Campbell, & Westen (2004)	Personality pathology in adolescents study (AACAP & APA clinicians)	USA	*N*=294 age range 14-18 52.9% female 85% Caucasian	Clinical sample (in treatment for PD pathology)	CBCL (clinician)	All	All ten DSM-IV PDs	Clinical diagnosis (DSM-IV Axis II criteria)	Cross-sectional
Ferguson, San Miguel, & Hartley (2009)	South Texas Hispanic cohort	USA	*N*=603 Mean (SD) age=12.4 (1.3) 48.8% female <4%Caucasian	Community sample	YSR, CBCL (parent)	Aggr.Behav. Delin. Beh.	ASPD	NLE (antisocial personality scale)	Cross-sectional
Fite, Greening, & Stoppelbein (2008)	Mississippi Inpatient cohort	USA	*N*=212 Mean (SD) age=8.3(2.4) 30% female 39% Caucasian	Inpatients	CBCL (parent or carer)	Aggr.Behav.	ASPD	APSD	Cross-sectional
Forsman, Larsson, Andershed,& Lichtenstein (2007)	Twin Study of Child and Adolescent Development (TCHAD)	Sweden	*N*=1,480 twin pairs age range 8-9 to 16-17 ?% female ?% Caucasian	Community cohort twin study	CBCL (parent)	Externalising	Psychopathy	YPI	Longitudinal (8 years follow-up)
Forsman, Lichtenstein, Andershed, & Larsson (2010)	Twin Study of Child and Adolescent Development (TCHAD)	Sweden	*N*=2,255 age range 8-9 to 16-17 ?% female ?% Caucasian	Community cohort twin study	YSR CBCL (parent)	Aggr.Behav. Delin. Beh.	Psychopathy	YPI	Longitudinal (8 years follow-up)

Goethals, Willigenburg, Buitelaar, & Van Marle (2008)	Nijmegen TBS (Dutch Entrustment Act) detainees	Netherlands	*N*=94 Mean (SD) age=38.8(8.3) 0% female 62% Caucasian	4 clinical groups: psychotic offenders with/without PD, non-psychotic offenders with a PD, & psychotic non-offenders without a PD	CBCL (actuarial data)	All	Cluster B	Clinical diagnosis (DSM-IV Axis II criteria)	Retrospective
Halperin, Rucklidge, Powers, Miller, & Newcorn (2011)	New York Children with disruptive behaviour disorders-biological correlates of ADHD study	USA	*N*=152 Mean (SD) age=9 (1.3) 13% female 22% Caucasian	Clinical: children with an ADHD diagnosis	YSR CBCL (parent)	CBCL-DP	Cluster A, B & C, Any PD	SCID-II	Longitudinal (9 years follow-up)
Kosson, Cyterski, Steuerwald, Neumann, & Walker-Mathews (2002)	North Carolina male delinquents	USA	*N*=115 Mean (SD) age=14.5(?) 0% female 28% Caucasian	Males on probation	CBCL (parent or carer)	Aggr.Behav. Delin. Beh.	Psychopathy	PCL: YV	Cross-sectional
Lexcen, Vincent, & Grisso (2004)	Massachusetts study	USA	*N*=481 Mean (SD) age=15.7(1), 0% female 38% Caucasian	Male youth involved with the juvenile justice system	YSR	Somatic C. Anxious/–Depressed Social Prob. Thought Pr. Attention Pr. Delin. Beh. Aggr. Behav.	Psychopathy	MACI - PCS	Cross-sectional
Meyer et al. (2009)	Maryland longitudinal study of children in families with and without maternal affective illness	USA	*N*=101 age range 5 -16 63% female 90% Caucasian	Children in families with and without maternal affective illness	CBCL (mother)	CBCL-DP	Cluster B	IPDE	Longitudinal (17 years follow-up)
Natsuaki, Cicchetti, & Rogosch (2009)	Minnesota longitudinal Summer Camp research programme	USA	*N*=174 age range 9-12 40% female ?% Caucasian	Maltreated and non-maltreated children of low SES	TRF (camp counsellors)	Externalising	Paranoid PD symptoms	OMNI-IV PD Inventory PPD scale	Longitudinal (3-6 years)

Penney, Moretti, & Da Silva (2008)	British Columbia behaviourally disordered youth	Canada	*N*=173 Mean (SD) age= 14.5 (1.7) 42% female 69% Caucasian	Clinic-referred adolescents (with serious behaviour disorder)	YSR	Aggr.Behav. Delin. Beh.	Psychopathy NPD	MACI-PCS MACI P-16 MACI egotistic personality scale	Cross-sectional
Rogosch & Cicchetti (2005)	Minnesota longitudinal Summer Camp research programme	USA	*N*=360 age range 6-12 49% female 17 % Caucasian	Maltreated and non-maltreated children of families of low SES	TRF (camp counselors)	All	BPD	BPD precursors composite	Cross-sectional
Sevecke, Lehmkuhl, & Krischer (2009)	Cologne-GAP Study (Gewalt=Violence; Aggression=Aggression; Persönlichkeit=Personality)	Germany	*N*=214 Mean (SD) age=17.7 (1.3) 58% female 81% Caucasian	Incarcerated juveniles	YSR	All	Psychopathy	PCL-YV	Cross-sectional
Sharp, Pane, et al. (2011)	Adolescent Treatment programme (ATP) at the Menninger Clinic	USA	*N*=111 Mean (SD) age=15.5 (1.4) 56% female ?% Caucasian	Adolescent inpatients	YSR	Internalising Externalising	BPD, ASPD	BPFSC APSD	Cross-sectional
Sharp, Ha, Michonski, Venta, & Carbone, 2012	Adolescent Treatment programme (ATP) at the Menninger Clinic	USA	*N*=190, Mean (SD) age=15.4 (1.5) 59% female 92% Caucasian	Adolescent inpatients	YSR CBCL (parent)	Internalising Externalising Total	BPD	BPFSC BPFSP PAI-A BOR CI-BPD	Cross-sectional
Sharp, Mosko, Chang, & Ha (2010)	Houston community cohort	USA	*N*=171 Mean (SD) age=13.5 (1.9) 0% female 62% Caucasian	Community sample of boys	YSR	Total & DSM-IV disorder specific scales	BPD	BPFSC BPFSP	Cross-sectional
Underwood, Beron, & Rosen (2011)	Texas longitudinal cohort	USA	*N*=255 age range 8-13 51% female 52% Caucasian	Community sample	TRF	Emotional (Carlson et al., 2009) Anxious/ Depressed Withdrawn/ Depressed Somatic C.	BPD NPD	IPDE-BPD subscale NPIC	Cross-sectional

Washburn, McMahon, King, Reinecke, & Silver (2004)	Chicago community cohort	USA	*N*=233 Mean (SD) age=12.5 (1.2) 63% female <2% Caucasian	Community sample	YSR	Anxious/ Depressed	NPD	NPI	Cross-sectional
Westen, Shedler, Durrett, Glass, & Martens (2003)	Personality pathology in adolescents study (AACAP & APA clinicians)	USA	*N*=296 age range 14-18 53% female 85% Caucasian	Clinical sample (in treatment for PD pathology)	CBCL (clinician)	All	All ten DSM-IV PDs	SWAP-200-A, Clinical diagnosis (DSM-IV Axis II criteria)	Cross-sectional
Wickline, Nowicki, Bollini, & Walker (2012)	Emory cohort	USA	*N*=65 Mean (SD) age=13.9(1.7) 35% female 85% Caucasian	Adolescents with SPD, other PDs & healthy controls	CBCL (parent)	Social & Thought Problems	SPD	SCID-II	Longitudinal (3 years)
Zelkowitz et al. (2007)	Montreal Child Psychiatry Day Hospital	Canada	*N*=59 adolescents Mean (SD) age=15.5(?) 19% female 82% Caucasian	Adolescents treated as children in a Child Psychiatry Day Hospital 5-7 years earlier	CBCL (parent)	All	BPD	CDIB-R	Longitudinal (5-7 years) PD assessment preceded ASEBA administration

Notes.
SCID-II: Structured Clinical Interview for DSM-III Axis II Personality Disorders (First, Gibbon, Spitzer, Williams, & Benjamin, 1997)
YSR: Youth Self-Report (Achenbach, 1991)
CBCL: Child Behaviour Checklist (Achenbach, 1991)
TRF: Teacher Report Form (Achenbach, 1991)
SIDP-IV: Structured Clinical Interview for DSM-IV Personality Disorders (Pfohl, Blum, & Zimmerman, 1997)
CDIB-R: Retrospective Diagnostic Interview for Borderlines (Greenman, Gunderson, Cane, & Saltzman, 1986)
BPFSC: Borderline Personality Features Scale for Children (Crick, Murray-Close, & Woods 2005)
BPFSP: Borderline Personality Features Scale for Parents (Sharp, Mosko, Chang, & Ha, 2011)
AACAP: American Academy of Child and Adolescent Psychiatry
APA: American Psychological Association
NPD: Narcissistic Personality Disorder
ASPD: Antisocial Personality Disorder
BPD: Borderline Personality Disorder
SPD: Schizotypal Personality Disorder

DSM: Diagnostic and Statistical Manual
CBCL-DP: Dysregulation Profile
ADHD: Attention Deficit Hyperactivity Disorder
SES: Socio-Economic Status
APSD: Antisocial Process Screening Device (Frick, Bodin, & Barry, 2000)
MACI: Millon Adolescent Clinical Inventory (Millon, 1993)
MACI –PCS MACI Psychopathy Content Scale (Murrie & Cornell, 2000)
MACI P-16 Millon Adolescent Clinical Inventory Psychopathy-16 (Salekin, Ziegler, Larrea, Anthony, & Bennett, 2003)
SWAP-200-A: Shedler-Westen Assessment Procedure-200 for Adolescents (Westen & Shedler, 1999a; 1999b)
OMNI-IV PD Inventory (Loranger, 2001)
IPDE-BPD: International Personality Disorder Examination-BPD (Loranger, 1995; Loranger, Janca, & Santorius, 1997)
NPI: Narcissistic Personality Inventory (Raskin & Hall, 1979)
NPIC: Narcissistic Personality Inventory–Children (Barry, Frick, & Killian, 2003)
DEP: Depressive disorders
NLE: Negative life events
YPI: Youth Psychopathic Traits Inventory (Andershed, Kerr, Stattin, & Levander, 2002)
PCL:YV: Psychopathy Check List: Youth Version (Forth, Kosson, & Hare, 2004)
PAI-A BOR: Personality Assessment Inventory for Adolescents (Morey, 2007)
CI-BPD: Childhood Interview for DSM-IV Borderline Personality Disorder (Zanarini, 2003)
Aggr.Behav.: Aggressive Behaviour
Delin. Beh.: Delinquent Behaviour
Somatic C.: Somatic Complaints

The 22 projects were carried out in the following countries across the world: USA (15), Germany (2), Canada (2), UK (1), Sweden (1), and the Netherlands (1). A total of 10,781 participants were included in the 29 studies, with considerable variation in the sample sizes (*M*=371, *SD*=674, Mdn=171, range 59 to 2960).

Twenty-three publications included clinical, forensic or high-risk samples recruited from a variety of services, and six of these studies also included comparison samples of healthy controls. The remaining six studies used community samples recruited through educational institutions or population registers. Studies took place in a variety of environments including schools, summer camps, prisons, forensic units, psychiatric emergency clinics, and both inpatient and outpatient settings.

Seventeen studies used the CBCL, fourteen the YSR, and four the TRF; six studies used both the CBCL and the YSR. In most studies, the CBCL was completed by caregivers. Different ASEBA scales were used in each study, and the majority of studies included in their results the two broadband scales (Internalising and Externalising Problems). A few studies used a specific combination of ASEBA items or scales.

All PD types were the subject of this review, and were used to organise the findings of the included studies. PD as an outcome was referred to in terms of diagnostic caseness (used categorically) across twelve studies. The remaining studies operationalised PD dimensionally (in terms of symptom level/severity). PD as an outcome was reported with regard to individual PD diagnoses, PD Clusters, and overall PD symptomatology. Most studies investigated BPD, or ASPD/psychopathy, and these will be the main focus of the review. The details of studies looking at two or more PDs will be presented in the order of first occurrence, whereas studies looking at all PD types, or any PD, will be presented at the end.

For the purposes of this review, a design was considered longitudinal when the administration of the YSR or CBCL or TRF preceded the assessment of PD. Eight studies

had a prospective design, and the length of time elapsed between ASEBA administration and PD assessment ranged from five to 17 years.

In the following sections, results of authors' univariate analyses are reported unless otherwise specified. In cases where effect sizes were not explicitly provided, they were calculated, where possible.

Cluster A PDs

Two studies focused on Cluster A personality pathology, paranoid PD (PPD) and schizotypal PD (STPD) in particular.

Paranoid PD

Natsuaki, Cicchetti, and Rogosch (2009) followed up children from a low socioeconomic background, some of whom were maltreated according to official records. The children joined an annual summer camp programme at least once between the ages of 9 to 12, and participated in a PD assessment three years later. Two to three camp counsellors rated the TRF, whereas the children completed the OMNI-IV Personality Disorder Inventory (Loranger, 2001) at follow-up. Because only five adolescents had *T* scores ≥ 70 on the OMNI-IV PPD scale (which indicates a clinically significant level of PD) participants were classified into three PPD groups as follows: low PPD ($T < 44$), moderate PPD ($44 < T < 54$) and high PPD ($T > 55$). Multilevel modelling analyses were then used to investigate whether the three groups had differed in their problem behaviours as children.

Results showed that children who developed high levels of PPD symptoms in adolescence had higher Externalising Problems in childhood, and interestingly had an upward growth in Externalising Problems between ages 9 to 12, whereas Externalising Problems of other groups slowly declined. This means that the discrepancy between the high and low PPD groups widened over time, ‘possibly forecasting the emerging PPD’ (Natsuaki et al., 2009, p.1191).

This is a noteworthy study for a number of reasons. Its prospective design and in particular the assessment of externalising problems at multiple time points extends the literature by suggesting that different developmental pathways may exist for children with PPD traits. However, without multiple measures of PPD, the authors were unable to examine the potential covariation of externalising and personality pathology across time. The multiple informant design is another strength of this study, as it reduced the potential for inflated shared-method variance.

However, the measurement of PPD relied on self-report alone and this is a limitation, as it is questionable whether individuals, especially adolescents, can accurately report their paranoid traits. Moreover, although inter-rater reliabilities for the TRF were satisfactory, according to Achenbach (1991) it is crucial that the respondent completing the TRF has known the young person for at least two months, and this is unlikely to have been the case in this study as the camp programme lasted just one week. In addition, the potential impact of the summer camp environment on children's behaviour is unknown. Furthermore, a low socioeconomic status (SES) sample was used, which limits the generalisability of the findings. Despite the above limitations, this study provided strong evidence for the predictive validity of the Externalising Problems scale (especially when used longitudinally).

Schizotypal PD

The three-year study by Wickline, Nowicki, Bollini, and Walker (2012) was part of a larger research programme looking at biological and behavioural aspects of STPD in adolescents. At Time 1, the sample consisted of 65 adolescents: some were diagnosed with STPD, some with other PDs (OPD), and some were non-psychiatric controls (NPC). Axis II diagnoses were based on the Structured Clinical Interview for DSM (SCID-II; First, Gibbon, Spitzer, Williams, & Benjamin, 1997), with high inter-rater reliability. About 70% of the sample agreed to participate in a follow-up assessment.

Analyses showed that at Time 1, STPD adolescents had more parent-rated Social and Thought Problems than both the NPC and the OPD groups. Furthermore, the OPD group had significantly more Thought Problems than the NPC group. At Time 2, the STPD group differed only from the NPC but not the OPD group (although differences were still in the predicted direction).

Given the small sample size, the results of this study cannot be easily generalised. In addition, a larger sample would have allowed the assessment of potential sex differences in the domains of Social and Thought Problems. Furthermore, PD was assessed only once, and it is unknown whether participants' diagnostic status at Time 1 remained the same at Time 2.

On the other hand, the use of parent reports and the assessment of PD with a structured diagnostic interview are significant strengths of this study. The inclusion of a second time point, at which the participation rate was quite high, provided further evidence for the criterion validity of the Social and Thought Problems scales. Results were largely replicated, yielding similar effect sizes.

Cluster B PDs

Antisocial PD (ASPD)-Psychopathy

Thirteen publications reported findings in relation to ASPD or psychopathy. These findings originated from eleven research projects; among these, two projects generated two publications each. Two citations stemmed from a longitudinal large-scale twin study, while the remaining studies were cross-sectional; among the latter studies, a few used the ASEBA to validate measures of juvenile psychopathy.

In a large longitudinal study of twin pairs born in Sweden, Forsman, Larsson, Andershed, and Lichterstein (2007) used the CBCL to assess parent-reported externalising behaviour problems when the twins were 8-9 and 13-14 years old.

Participants that scored above the 75th percentile at both 8-9 and 13-14 years were defined as having persistent externalising behaviour problems. Self-reports of personality constellation were obtained when participants were 16-17 years old with the Youth Psychopathic traits Inventory (YPI; Andersheld, Kerr, Stattin, & Levander, 2002). The YPI consists of ten subscales and has three higher order factors, namely the interpersonal (grandiose/manipulative), affective (callous/unemotional) and lifestyle factor (impulsive/ irresponsible).

The correlations between the CBCL Externalising Problems and the YPI total and subscale scores were overall statistically significant (given the large sample size) but very modest, providing weak evidence of predictive validity. The associations were somewhat higher for externalising behaviour at age 13-14, as compared to ages 8-9, probably resulting from the briefer period that elapsed between the two assessment points. However, when persistent externalising behaviour was used in the analysis, a moderating effect of gender was detected: Compared to a male control group, boys with persistent externalising behaviour scored higher on the YPI, and its callous/unemotional and impulsive/irresponsible dimensions. However, these effects did not apply to girls; this may suggest that the measures used (of persistent externalising behaviour problems, or of psychopathy) may have been inappropriate for girls.

This study has significant strengths, including the multiple-source assessment, a longitudinal design with a relatively high response rate and a large sample size, which permitted a gender-specific analysis of the findings. At the same time, the results of this study are limited by the non-clinical CBCL cutoffs used, which resulted in high prevalence rates of persistent externalising problems. Furthermore, psychopathic personality was measured on a continuous self-report scale; a more robust operationalisation would have included collateral reports.

A more recent citation (Forsman, Lichterstein, Andershed, & Larsson, 2010) stemming from the same study followed a combined informant approach that used both the CBCL and the YSR Delinquent and Aggressive Behaviour scales to measure antisocial behaviour. In this study, persistent antisocial behaviour was measured by summing parent-reports at age 8-9 with combined self- and parent-reports at age 13-14 and 16-17 years. Five items that were not included in both the self- and parent-reports were removed, and four items that were very similar to YPI items were also excluded to avoid spurious associations due to item overlap.

Unlike the previous study, correlations between the YPI (Andersheld et al., 2002) total score and antisocial behaviour were substantial and similar for males and females at both time points; the cross-lagged correlation between antisocial behaviour at Time 1 and psychopathic personality at Time 2 were significant in both genders, in support of the predictive validity of the CBCL. However, structural equation modelling showed that psychopathic personality in mid-adolescence predicted antisocial behaviour in adulthood, but not the other way around. When the authors' measure of persistent antisocial behaviour (from age 8-9 to age 16-17) was used, it was found that it explained less than 1% of the total variance in psychopathic personality at Time 2.

This study extends the findings of the previous publication of Forsman et al. (2007) by using a more sophisticated measure of persistent antisocial behaviour and more advanced methods of statistical analysis; this allowed the exploration of longitudinal associations. Overall, the predictive validity of the ASEBA externalising scales was weak in this study, but increased when ASEBA was administered more than once.

The only UK study reviewed here was conducted by Dolan and Rennie (2006), who used the Psychopathy Check List: Youth Version (PCL:YV; Forth, Kosson, & Hare, 2004) to test 115 male, incarcerated adolescents with conduct disorder. The PCL:YV is a multi-item symptom construct rating scale with an expert-rater format. Psychopathy ratings are

based on a semi-structured interview, a review of case history information and behavioural observation cross-checked with collaterals. The PCL:YV was designed to measure interpersonal and affective characteristics as well as overt behaviours related to adolescent psychopathy. Factor analyses of this scale have resulted in a two-factor model (reflecting interpersonal/affective dimensions and behavioural/lifestyle features of psychopathy), or a three-factor model (reflecting interpersonal, affective and behavioural/lifestyle features).

It was found that the PCL:YV total and two-factor scale scores correlated positively with Delinquent and Aggressive Behaviour and Attention and Externalising Problems on the CBCL (completed by carers). The analysis in relation to a three-factor model indicated that the concurrent validity was strongest in relation to the third (behavioural/lifestyle features) factor. There were no other significant correlations between the PCL:YV and the CBCL Internalising broadband or narrowband scales.

Overall, the results of this study support the concurrent validity of the CBCL, and the evidence is strengthened by the multiple informant design and the fact that psychopathy ratings were given by trained observers. On the other hand, the cross-sectional design is a limitation, and the generalisability of the findings in healthy controls and females is questionable, in light of the moderating effect of gender which has been elsewhere detected (e.g. Forsman et al., 2007).

In another citation of the same study, Dolan and Rennie (2007) used the YPI (Andersheld et al., 2002) to measure psychopathy. Significant but modest positive correlations were reported between CBCL Attention Problems and Externalising scales and most YPI factors, in line with the findings of Forsman et al. (2007). On the other hand, the associations with Internalising Problems were close to zero, as were the associations reported when the PCL:YV was used. These results are consistent with the ones reported

by the authors when they used the PCL:YV (Forth et al., 2004) and the evidence is of a similar quality to their 2006 publication.

In another study using the PCL:YV (Forth et al., 2004), Kosson, Cyterski, Steuerwald, Neumann, and Walker-Matthews (2002) assessed 115 adolescent males on probation, whose parents and guardians completed the CBCL Delinquent and Aggressive Behaviour, and the Attention Problems scales. To ensure that correlations between the CBCL and PCL:YV scores were not inflated by collaterals that provided similar information in the interview and on CBCL items, PCL:YV items likely to be scored on the basis of information relevant to each CBCL scale were removed.

Corrected PCL:YV scores correlated significantly with the Delinquent and Aggressive Behaviour scales of the CBCL, but did not correlate with Attention Problems scores, in contrast with the findings reported by Dolan and Rennie (2006). Differences in the ethnic mix of the samples may account for these discrepant findings. In addition, the removal of PCL:YV items that overlapped with the CBCL (in order to minimise the problem of overlapping item content) may have reduced the concurrent validity estimates reported. Other limitations include the cross-sectional design and the specific characteristics of the sample, which reduce the external validity of the findings.

Sevecke, Lehmkuhl, and Krischer (2009) also used the PCL:YV (Forth et al., 2004) to assess a group of adolescent male and female detainees in Germany. Participants completed the YSR, and cross-sectional linear regressions were conducted to explore which factors predicted the PCL:YV total score and its four dimensions (i.e. interpersonal, affective, lifestyle, and antisocial features), in accordance with Hare's 4-factor model (Vitacco, Neumann, & Jackson, 2005).

For males, evidence of concurrent validity was reported for the Aggressive Problems scale, while the Anxious/Depressed scale was associated negatively with

psychopathy. Among females, psychopathy was also predicted by Aggressive Behaviour, but the Anxious/Depressed Problems scale was not a significant predictor of the PCL:YV total score.

By including female and male delinquent adolescents, this study demonstrated that the absence of nervousness (operationalised using the YSR Anxious/Depressed scale) characterises males with psychopathic features, but not females. Furthermore, the use of the PCL:YV addressed the issue of shared method variance. On the other hand, the main limitation of this study is its cross-sectional design and the lack of clarity in the presentation of findings.

Two studies used the YSR to validate the self-report Psychopathy Content Scale (PCS; Murrie & Cornell, 2000) derived from the Millon Adolescent Clinical Inventory (MACI; Millon, 1993). The PCS was developed by selecting 25 (out of the 160) items from the MACI that were theoretically related to psychopathy and then removing five items that decreased the internal consistency of the scale, which resulted in a 20-item, single-factor scale.

In a large sample of adolescent boys, Lexcen, Vincent, and Grisso (2004) used a two-factor solution of the PCS. Factor 1 appeared to be consistent with the interpersonal and affective dimensions of psychopathy, whereas factor 2 related to the antisocial, lifestyle-type characteristics of psychopathy.

It was found that both PCS factors correlated positively with all YSR scales, with the exception of the Social Problems and Anxious/Depressed scales. To examine differences between high and low PCS scores, cutoff scores were determined so that approximately 30% of participants were identified as high scorers; YSR scorers were also dichotomised to identify clinically significant scores. Greater proportions of high PCS factor 1 scorers fell above cutoff on the Thought Problems, Delinquent and Aggressive Behaviour, and likewise significantly greater proportions of high PCS factor 2 scorers were

above cutoff on all scales except for Somatic Complaints and Social Problems. Following regression analyses, it was found that Delinquent Behaviour and Thought Problems were positive predictors of both PCS factors, whereas the Anxious/Depressed and Social Problems scales were negative predictors. The negative correlation with Social Problems is of interest and may reflect the interpersonal confidence and grandiosity that is known to characterise psychopathy.

Although this study used a large sample size, it is limited as it included only male participants. Another limitation is the exclusive reliance on self-report measures, none of which includes a measure of response bias (Lexcen et al., 2004). This may have inflated the concurrent validity estimates reported due to shared method variance.

Penney, Moretti, and Da Silva (2008) also used the Aggressive and Delinquent Behaviour scales of the YSR to validate the MACI Psychopathy scales in a sample of clinic-referred adolescents with serious behavioural disorders. Along with the PCS (Murrie & Cornell, 2000) which was also used in the previous study, the Psychopathy-16 (P-16) Scale of the MACI was administered (Salekin, Ziegler, Larrea, Anthony, & Bennett, 2003). The P-16 was developed by selecting 25 MACI items that conceptually mapped onto the PCL-R (Hare, 1998) and removing nine items that decreased the internal consistency of the scale. This resulted in a scale that more explicitly outlines the interpersonal and affective features of the construct (Cooke & Michie, 2001). In this study, a *T* score of 65 was used to dichotomise YSR scores that fell above and below a clinically significant threshold. An exploratory factor analysis of the P-16 in this study found three factors, namely a) lack of empathy and callousness, b) egocentricity (a conceited and manipulative style), and c) antisociality (antisocial and law-breaking behaviours).

In keeping with the findings of Lexcen et al. (2004), results showed that both factors on the PCS and all three factors on the P-16 correlated with Aggressive and Delinquent Behaviour on the YSR. The relatively extreme characteristics of the sample,

the reliance of self-report measures and the cross-sectional design used are obvious methodological limitations.

In a study of psychopathic traits among younger children, Fite, Greening and Stoppelbein (2008) tested 6-12 year-old psychiatric inpatients. Their parents or guardians completed the Aggressive Problems scale of the CBCL, as well as the 20-item Antisocial Process Screening Device (APSD; Frick, Bodin, & Barry, 2000), which is the only measure of psychopathic-like traits known for young children of this age. A total score and three subscale scores are derived: callous/unemotional traits, narcissism, and impulsivity.

Correlations with the CBCL Aggressive Behaviour scale were highest for the narcissism subscale and the APSD total score, but were also significant for the callous/unemotional and impulsivity subscales. Moreover, Aggressive Behaviour accounted for a substantial proportion of the variance of all three APSD subscales, as well as the total score. These results support the concurrent validity of the CBCL Aggressive Problems, but the estimate reported may have been inflated due to shared method variance. Furthermore, the cross-sectional design is a limitation, and the generalisability of the findings beyond an inpatient sample requires further investigation. At the same time, the recruitment of an understudied population (inpatient children) is a definite strength of this study.

In another study using the APSD that was conducted with inpatient adolescents, Sharp, Pane, et al. (2011) found that the APSD was positively correlated with the YSR Internalising and Externalising Problems, and compared to Internalising Problems, the association with Externalising Problems was significantly stronger. While the recruitment of a less well studied population is a significant strength, this study is limited by its cross-sectional design, the shared method variance and the reliance on a self-report questionnaire to measure psychopathic traits.

The remaining two studies focused on ASPD features rather than psychopathy. In the first study, Ferguson, Miguel, and Hartley (2009) tested 603 predominantly Hispanic children (aged 10-14 years) using the antisocial personality scale from the Negative Life Events instrument (NLE; Paternoster & Mazerolle, 2004). They also obtained self and parent ratings of Aggressive and Delinquent Behaviour using the YSR and the CBCL respectively. Correlations were positive with both externalising scales, and compared to the CBCL, the effect sizes were somewhat larger for the YSR. This seems to be the result of shared method variance, as both the NLE and the YSR are self-report measures. Beyond its cross-sectional design, another limitation of this study is the Hispanic-majority sample, which limits the external validity of the findings.

In another community study in Germany, Barnow, Lucht, and Freyberger (2005) interviewed adolescents older than 15 years of age using the SCID-II (First et al., 1997). A value of more than three in the self-rating section of the interview indicates a tendency toward ASPD. Participants completed the YSR, and their mothers completed the CBCL. The correlation analyses revealed significant relationships between ASPD and Delinquent and Aggressive Behaviour for combined YSR and CBCL ratings.

Limitations of this study include its cross-sectional design, and the availability for data analysis of only one section of the originally identified sample (i.e. adolescents older than 15 years of age). Furthermore, the authors recognised that the observations could not be considered fully independent because some children came from the same family. The use of a clinical interview to measure ASPD and the use of multiple informants (parent and self-report) are important strengths of this study. However, concurrent validity estimates were not reported separately for the YSR and the CBCL.

Table 3 presents a summary of the findings reviewed so far.

Table 3. *Criterion validity estimates concerning the association between ASPD/psychopathy and ASEBA Delinquent and Aggressive Behaviour, Internalising, Externalising and Total Problems scales*

Study	**Delinq. Behav. effect size *r* (*d*)**	**Aggr. Behav. effect size *r* (*d*)**	**Internal. Probl. effect size *r*(*d*)**	**External. Probl. effect size *r* (*d*)**	**Total Probl. effect size *r*(*d*)**	**ASEBA forms used**	**PD criterion used**	**Shared method variance**	**Validity Type**	**If group comparisons, non-PD group type**
Barnow et al.(2005)	.43**	.34**				CBCL, YSR	SCID-II	NO	Concurrent	
Dolan & Rennie (2006)	.37***	.37***	.002	.39***		CBCL	PCL:YV 2-factor model	NO	Concurrent	
	.35***	.31**	-.01	.34***		CBCL	PCL:YV 3-factor model	NO	Concurrent	
Dolan & Rennie (2007)	.20*	.21*	.00	.21*	.20*	CBCL	YPI	NO	Concurrent	
	.30* (*d*=.62)	.33** (*d*=.70)	.23 (*d*=.48)	.34** (*d*=.72)	.32* (*d*=.67)	CBCL	YPI (dich.)	NO	Concurrent	Clinical group with conduct disorder, non-psychopathic like
Dutra et al.[a](2004)	.81***	.77***	-.06***	.84***	.55***	CBCL	DSM-IV composite score	YES	Concurrent	
Ferguson et al. (2009)	.42**	.35**				YSR	NLE (antisocial personality scale)	YES	Concurrent	
	.28**	.27**				CBCL	NLE (antisocial personality scale)	NO	Concurrent	
Fite et al. (2008)		.69 (*p* not Reported)				CBCL	APSD	YES	Concurrent	

Forsman et al. (2007)				.13**	CBCL	YPI	NO	Concurrent	
				.21***	CBCL	YPI	NO	Predictive	
				.20** (*d*=.40)	CBCL Persistent external. (dich.)	YPI	NO	Concurrent	Community cohort without persistent ext. behaviour
Forsman et al. (2010)	.52***				CBCL & YSR	YPI	YES (partly)	Concurrent	
	.43***				CBCL & YSR	YPI	YES (partly)	Concurrent	
	.30***				CBCL & YSR	YPI	YES (partly)	Predictive	
	.36***				CBCL & YSR	YPI	YES (partly)		
Kosson et al. (2002)	.47***	.40***			CBCL	PCL:YV	NO	Concurrent	
Lexcen et al. (2004)	.65 (sign.at Bonferroni corrected error rate)	.62 (sign.at Bonferroni corrected error rate)			YSR	MACI PCS factor 1	YES	Concurrent	
	.66 (sign.at Bonferroni corrected error rate)	.49 (sign.at Bonferroni corrected error rate)			YSR	MACI PCS factor 2	YES	Concurrent	
Penney et al. (2008)	.59***	.40***			YSR	MACI PCS total	YES	Concurrent	
	.51***	.40***			YSR	MACI P-16	YES	Concurrent	
Sharp, Pane et al. (2011)			.26*	.61**	YSR	APSD	YES	Concurrent	
Westen et al.[a] (2003)	.76***	.73***	-.06***	.80***	CBCL	DSM criteria	YES	Concurrent	
	.69***	.80***	-.26***	.81***	CBCL	SWAP-200 A	YES	Concurrent	

Notes. *$p < .05$ **$p < .01$ *** $p < .001$
a. These studies are reviewed in the section: “All PDs”, p. 57.
ASPD: Antisocial Personality Disorder
CBCL: Child Behaviour Checklist (Achenbach, 1991)
YSR: Youth Self-Report (Achenbach, 1991)
SCID-II: Structured Clinical Interview for DSM-III Axis II Personality Disorders (First, Gibbon, Spitzer, Williams & Benjamin, 1997)
PCL:YV: Psychopathy Check List: Youth Version (Forth, Kosson, & Hare, 2004)
YPI: Youth Psychopathic Traits Inventory (Andershed, Kerr, Stattin & Levander, 2002)
dich.: dichotomised
NLE: Negative life events
APSD: Antisocial Process Screening Device (Frick, Bodin, & Barry, 2000)
Sign.: significant
Psychop.: psychopathy
Del. beh.: delinquent behaviour
MACI: Millon Adolescent Clinical Inventory (Millon, 1993)
MACI-PCS MACI Psychopathy Content Scale (Murrie & Cornell, 2000)
MACI P-16 Millon Adolescent Clinical Inventory Psychopathy-16 (Salekin, Ziegler, Larrea, Anthony, & Bennett, 2003)
SWAP-200-A: Shedler-Westen Assessment Procedure-200 for Adolescents (Westen & Shedler, 1999a; 1999b)

As Table 3 shows, there was a wide variation in the size of the reported correlations. For instance, the correlation with Externalising Problems ranged from .13 to .84. On the whole, the direction of the effect was positive for both Delinquent and Aggressive Behaviour scales and the Externalising Problems broadband scale. However, for Internalising Problems, findings were rather inconsistent, with some studies reporting positive associations, some negative associations but other studies finding no evidence of criterion validity.

Borderline PD (BPD)

Nine studies used the ASEBA to investigate BPD, the most widely researched and written about single-PD domain in adolescents, as in adults. Two of these studies had a longitudinal design (i.e. ASEBA administration preceded PD assessment).

In a prospective study in Germany, Arens, Grabe, Spitzer, and Barnow (2011) examined whether internalising and externalising problems measured during adolescence with the YSR and the CBCL contributed to the risk of BPD, diagnosed (with satisfactory inter-rater agreement) using the SCID-II (First et al., 1997) five years later.

They found that young adults diagnosed with BPD had higher Internalising and Externalising Problems in adolescence, compared not only to age, sex and educational level-matched healthy controls, but also compared to a matched group of participants diagnosed with depressive disorders. It is of interest that the predictive validity estimates tended to be larger for the YSR, compared to the CBCL ratings. Arens et al. (2011) also found that Internalising (but not Externalising) Problems predicted the risk of BPD diagnosis vs. no disorder, and an increased level of Internalising Problems was the only significant predictor that distinguished between BPD diagnosis and depressive disorders. However, it was not clarified whether these predictors were based on the CBCL or the YSR, or whether a composite measure was derived using both scales.

Strengths of this study include its longitudinal design, the inclusion of both self-and parental reports, the use of a well-established diagnostic interview to assess BPD, and the inclusion of a clinical control group, which addressed the question of the specificity of the ASEBA to BPD. The unbalanced sex ratio and the small sample size are limitations of the study.

In another longitudinal study of development and adaptation in a poverty sample of young mothers and their first-born children, Carlson, Egeland, and Sroufe (2009) examined the antecedents and developmental course of BPD symptoms prospectively from infancy to adulthood. When children were aged 12, the TRF was completed by the child's teacher. When participants were 28 years old, BPD symptom counts (ranging from 0 to 7) were derived from the SCID-II (First et al., 1997).

For the purposes of this study, to represent instability and disturbance in emotional, behavioural, attentional, and relational domains, seven items were selected from the TRF for each domain to represent core self processes underlying characteristics of borderline personality pathology (Geiger & Crick, 2001)[1]. All scales had adequate internal consistency (alpha ranged from .77 to .87), and there was evidence of predictive validity: Correlational analyses confirmed moderate associations between adult borderline symptoms and disturbance across these four domains of functioning in middle childhood/early adolescence. Furthermore, when entered simultaneously as predictor variables of BPD in combination with measures of self-representation and parent-child relationship, the Emotional Instability subscale demonstrated a marginally significant influence.

[1] Emotional items included: "cries a lot", "nervous high-strung or tense", "stubborn, sullen, or irritable", "sudden changes in mood", "sulks a lot", "temper tantrums or hot temper", and "unhappy, sad, or depressed" (α =.77). Behavioural items included: "impulsive, acts without thinking", "fails to finish things", "destroys own things", "accident prone, gets hurt a lot", "behaves irresponsibly", explosive and unpredictable behaviour", and "easily frustrated" (α = .79). Attentional items included: "can't concentrate, can't pay attention", "confused or seems to be in a fog", "daydreams or gets lost in thoughts", "absorbed with picking at skin/body", "sleeps in class", "stares blankly", and "inattentive, distracted" (α = .81). Relational items included: "bullying, meanness to others", "destroys property of others", "disturbs others", "doesn't get along with others", "gets in fights", "physically attacks people", and "threatens people" (α = .87).

This is an outstanding study for various reasons. The length of the follow-up (from infancy to adulthood) allowed the exploration of the long-term predictive validity of the TRF. Furthermore, the use of specific items from the TRF to represent instability and disturbance in various domains is innovative, and although these subscales are less well established (compared to the problem scales of the CBCL), the usefulness of these items for assessing emerging PD is worthy of further exploration. The use of multiple independent assessment methods and informants, including the SCID, is another strength of this study. Furthermore, the use of a dimensional analytic approach allowed the examination of borderline phenomena in a community sample; however, the generalisability of these findings to clinical settings remains to be established.

In a five-year prospective study, in which ASEBA and PD measures were administered concurrently, Underwood, Beron, and Rosen (2011) examined in a community sample of children the relationship between adjustment problems at age 14 and developmental trajectories of social and physical aggression. At age 14, participants completed the BPD subscale of the International Personality Disorder Examination (IPDE; Loranger, 1995; Loranger, Sartorius, & Janca, 1997), while their eighth-grade teachers rated them on the Anxious/Depressed, Delinquent, Withdrawn/Depressed and Somatic Problems syndromes of the TRF. To explore problems in emotion regulation, the Emotional Instability subscale (developed by Carlson et al., 2009 and described previously) was also used.

Results showed that surprisingly, apart from some positive trends, no subscale (including the Emotional Instability subscale) correlated significantly with BPD features; this contradicts the findings of most other studies reviewed. This finding was not discussed by the authors, probably because the focus of the study was rather different (the developmental trajectories of aggression). It is possible that measurement issues (the reliability of the borderline personality features measure was in the questionable range) as well as sampling issues (such as the low participation rate) may have contributed to this finding. On the other hand, the use of a quite large, typically developing sample is a methodological advantage.

In another prospective study examining the long-term outcome of borderline pathology of childhood (BPC), established using the child version of the Retrospective Diagnostic Interview for Borderlines (CDIB-R; Greenman, Gunderson, Cane, & Saltzman, 1986) Zelkowitz et al. (2007) followed a sample of children with BPC from middle childhood into adolescence, and compared them with a group of former child psychiatry patients with no history of BPC. As adolescents, participants were reassessed for BPD with the equivalent diagnostic interview (DIB-R; Zanarini, Gunderson, Frankenburg, & Chauncey, 1989). However, among the 24 adolescents with BPC, only four met criteria for a current diagnosis of BPD, and among the adolescents with no history of BPC, one participant met criteria for BPD.

It was found that the group with a history of BPC exhibited more Thought Problems and Aggressive Behaviour, and scored higher on the Withdrawn/Depressed and Anxious/Depressed syndromes. There was also evidence of increased risk for Internalising Problems.

For the purposes of this review, the main limitation of this study is that no comparisons were made between participants with and without current BPD identified in their CBCL scores. Moreover, the temporal stability between borderline pathology in childhood and borderline pathology in adolescence was low; therefore the reported comparisons are of limited usefulness. Another limitation of this study includes the high sample attrition. Furthermore, this study focused on a quite extreme group of children at risk for later PD.

In another publication based on the summer camp project for low SES children mentioned above, Rogosch and Cicchetti (2005) investigated the potential precursors to BPD. As the purpose of this study was the detection of people who are *vulnerable* to developing BPD later in life, rather than the identification of children with borderline pathology, a BPD precursors composite was created by combining self-report, peer-report, and counsellor-report measures assessing personality features, representational models of

self, relationship difficulties with peers and adults, and suicidal/self-harm behaviour. The TRF item concerning self-harm and suicidal behaviour as observed by counsellors was also included in the composite.

It was found that the BPD precursors composite total score correlated with both internalising and externalising pathology on the TRF, and the concurrent validity of Externalising Problems was stronger. The BPD composite correlated significantly with all narrowband scales and strongest associations were evidenced for Aggressive Behaviour, Social Problems, Delinquent Behaviour, and Attention Problems.

The use of multiple developmental constructs to derive a BPD composite is a strength of this study as a whole, but for the purposes of this review, the inclusion of the TRF self-harm item may have inflated some of the associations observed due to item overlap. Furthermore, and in contrast to Natsuaki et al.'s (2009) study of PPD, this study is limited by its cross-sectional design and the absence of a criterion used to empirically validate the BPD precursors composite. On balance, this study presents only limited evidence for the concurrent validity of the TRF.

In a study validating the Borderline Personality Features Scale for Children (BPFSC; Crick, et al., 2005) Sharp, Mosko, Chang, and Ha (2011) administered the BPFSC and the BPFSP (parental equivalent) to a community sample of 8 to 18-year-old boys. Given the non-clinical nature of the sample, the BPFS data were positively skewed and the 80th percentile was therefore used as a cutoff to create high- and low-scoring groups on the BPFSC and BPFSP. The CBCL and the YSR were also administered. Total Problems with *T* scores > 65 was used as a cutoff to discriminate between clinical and non-clinical populations as indicated by Achenbach and Rescorla (2001). Moreover, a few DSM-IV Axis-I oriented YSR/CBCL scales were included in the analysis (i.e. affective problems, anxiety problems, somatic problems, attention deficit/hyperactivity problems, oppositional defiant problems, and conduct problems) to identify disorder-specific problems among young people.

Within informant type, the BPFSC and the BPFSP were positively and strongly correlated with Total Problems on the YSR and the CBCL respectively. These associations remained significant but their strength decreased substantially when different informants rated the BPFS and the ASEBA. Furthermore, the high-BPD trait group (based on both parents' and self-reports) had significantly more CBCL and YSR Total Problems respectively and significantly more symptoms for all corresponding ASEBA DSM-oriented scales. Across informants, however, the associations between the BPFSC and BPFSP with the YSR and CBCL DSM-oriented problems were quite modest. Compared to Total or Internalising Problems, Externalising Problems had a stronger association with the BPFSC and BPFSP.

This study was one of the few to allow exploration of the effect of shared method variance. The correlations between ASEBA measures and the BPFS were considerably higher within informant type rather than across informants, which needs to be taken into account when evaluating the criterion validity of the ASEBA. Another strength of this study is the recruitment of a community sample of boys, given the over-focus on females in studies of BPD. However, the generalisability of the findings to females or other populations remains to be tested.

A limitation of the study is that the BPFS was used without empirically established clinical cutoffs, and a rather arbitrary cutoff point was used instead. Moreover, although a dimensional approach to measuring BPD has advantages in a community sample, there is only one study that has validated the BPFS against structure-interview diagnosis (Sharp, Ha, Michonski, Venta, & Carbone, 2012). Therefore, more research is needed to investigate the psychometric properties of the ASEBA.

In another study using the BPFSC (Crick et al., 2005), Sharp, Pane, et al. (2011) tested 111 adolescent inpatients and found that the BPFSC was positively correlated with the YSR Internalising and Externalising Problems. The magnitude of the associations was comparable to the correlations reported between the BPFSC and the YSR Total Problems score in the previous study. Compared to Internalising Problems, the correlation between the

BPFSC and Externalising Problems was somewhat stronger. Overall, these findings replicate the results of the previous study, but this time with a psychiatric sample. Limitations of this study include its cross-sectional design, the shared method variance and the reliance on a self-report questionnaire to measure BPD.

In the same adolescent inpatient setting, Sharp et al. (2012) used the YSR and the CBCL together with a clinical interview (CI-BPD; Zanarini, 2003) and two questionnaire-based measures of adolescent BPD, namely the BPFSC (Crick et al., 2005) and the Personality Assessment Inventory-Borderline Features scale (PAI-BOR; Morey, 2007). They found that adolescents with a BPD diagnosis scored significantly higher on both the YSR and the CBCL Internalising, Externalising and Total Problems. The effect sizes for the self-reported data were somewhat larger.

Furthermore, the Internalising and Externalising Problems scales of the YSR and the CBCL correlated significantly with both the interview measure of BPD, and the two self-report measures, with higher associations reported for the interview measure (CI-BPD) and the self-report measures (BPFSC and PAI-BOR), compared to parent-report data (BPFSP).

Although this is a cross-sectional study, the robustness of these findings lies in the use of multiple measures of BPD, namely the inclusion of both self and parent reports, as well as a diagnostic interview. Furthermore, the large inpatient sample allowed comparisons between the BPD and a clinical control group with psychopathology of comparable severity. Like Arens et al.'s (2011) study above, this provides some evidence for the specificity of ASEBA in relation to BPD. On the other hand, the study is limited by sampling issues; i.e. participants were predominantly White adolescents of high SES, at the severe end of the psychopathology spectrum who failed to respond to previous treatments. It is unclear if these findings would apply to other groups of adolescents, and further research with community samples would be needed to maximise generalisations.

Finally, Burnette and Reppucci (2009) assessed 121 incarcerated, teenage girls using the Structured Interview for DSM-IV Personality (SIDP-IV; Pfohl, Blum, & Zimmerman, 1997). Interview items were scored on a 0-2 scale and the sum across the nine criteria assessing BPD symptoms was used as a continuous measure of BPD. Excellent inter-rater reliability was established using paired ratings of nine cases. Participants also completed the YSR Aggressive Behaviour scale, which correlated moderately with BPD.

This study is limited by its cross-sectional design, and the evidence of concurrent validity reported may be specific to this group of adolescents that demonstrated deficits in multiple areas of functioning and constitute the most extreme end of the spectrum. Furthermore, the concurrent validity of the remaining YSR scales was not investigated. On the other hand, the use of an interview with high diagnostic reliability among interviewers is an obvious advantage.

Tables 4 and 5 present a summary of the above findings.

Table 4. *Effect sizes on the relationship between BPD and ASEBA's narrowband syndromes*

Study	**Withdrawn/ Depressed effect size *r* (*d*)**	**Somatic Complaints effect size *r* (*d*)**	**Anxious/ Depressed effect size *r* (*d*)**	**Social Problems effect size *r* (*d*)**	**Thought Problems effect size *r* (*d*)**	**Attention Problems effect size *r* (*d*)**	**Delinquent Behaviour effect size *r* (*d*)**	**Aggressive Behaviour effect size *r* (*d*)**	**PD criterion used**
Burnette et al. (2007)	.24**	.26**	.49**				.33**	.31**	SIDP-IV Vulnerable factor
Burnette & Reppucci (2009)								.37**	SIDP-IV BPD
Dutra et al. (2004)[a]	.25**	.13	.37**	.30**	.22**	.35**	.34**	.45**	DSM-IV composite score
Rogosch & Cicchetti (2005)	.15**	.14**	.30***	.63***	.29***	.47***	.52***	.72***	BPD precursors composite
Underwood et al.(2011)	.14	-.02	.02				.13		IPDE BPD
Westen et al. (2003)[a]	.23***		.35***				.35***	.45***	DSM-IV criteria
	.14		.33***				.33***	.55***	SWAP-200 A
Zelkowitz et al. (2007)[b]	.33* (*d*=.72)	.21 (*d*=.42)	.30* (*d*=.63)	.14 (*d*=.29)	.28* (*d*=.59)	.23 (*d*=.48)	.25 (*d*=.52)	.28* (*d*=.57)	CDIB-R

*$p < .05$ **$p < .01$ *** $p < .001$

Notes. The TRF was used in Underwood et al.'s (2011) and Rogosch & Cicchetti's (2005) studies, and the YSR was used in Burnette et al.'s (2007) and Burnette & Reppucci's studies (2009).The CBCL was used in the remaining studies. All reported associations are cross-sectional (concurrent validity), apart from Zelkowitz et al.'s (2007) study, in which PD assessment preceded ASEBA administration

a: Shared method variance. These studies are reviewed in the section: "All PDs", p. 57.

b: The comparisons made were between a BPD and a non-BPD clinical control group

SIDP-IV: Structured Clinical Interview for DSM-IV Personality Disorders (Pfohl, Blum, & Zimmerman, 1995)

IPDE-BPD: International Personality Disorder Examination-BPD (Loranger, 1995; Loranger, Janca, & Santorius, 1997)

CI-BPD: Childhood Interview for DSM-IV Borderline Personality Disorder (Zanarini, 2003)

SWAP-200-A: Shedler-Westen Assessment Procedure-200 for Adolescents (Westen & Shedler, 1999a; 1999b)

CDIB-R: Retrospective Diagnostic Interview for Borderlines (Greenman, Gunderson, Cane, & Saltzman, 1986)

Table 5. *Effect sizes on the relationship between BPD and ASEBA Internalising, Externalising and Total Problems scales*

Study	**Internalising Problems effect size *r* (*d*)**	**Externalising Problems effect size *r* (*d*)**	**Total Problems Effect size *r* (*d*)**	**ASEBA forms**	**PD criterion used**	**Shared method variance**	**Validity type**	**If group comparisons, non-PD group type**
Arens et al. (2011)	.39** (*d*=.84)	.41* (*d*=.888)		YSR	SCID-II (dich.)	NO	Predictive	Depressive group
	.74*** (*d*=2.21)	.64*** (*d*=1.69)		YSR	SCID-II (dich.)	NO	Predictive	Healthy controls
	.38** (*d*=.83)	.48** (*d*=1.10)		CBCL	SCID-II (dich.)	NO	Predictive	Depressive group
	.60*** (*d*=1.52)	.53*** (*d*=1.26)		CBCL	SCID-II (dich.)	NO	Predictive	Healthy controls
Burnette et al. (2007)	.42**	.33**		YSR	SIDP-IV Vulnerable factor	NO	Concurrent	
Dutra et al. (2004)[a]	.33***	.44***	.53***	CBCL	DSM-IV composite score	YES	Concurrent	
Rogosch & Cicchetti (2005)	.25***	.71***	.70***	TRF	BPD precursors composite	NO	Concurrent	
Sharp, Mosko, et al. (2011)			.63**	YSR	BPFSC	YES	Concurrent	
			.64**	CBCL	BPFSP	YES	Concurrent	
			.33**	YSR	BPFSP	NO	Concurrent	
			.21*	CBCL	BPFSC	NO	Concurrent	
Sharp, Pane, et al. (2011)	.53**	.60**		YSR	BPFSC	YES	Concurrent	

Sharp et al. (2012)	.42***	.54***		YSR	CI-BPD	NO	Concurrent	
	.29*** (*d*=.60)	.44*** (*d*=.98)	46*** (*d*=1.03)	YSR	CI-BPD (dich.)	NO	Concurrent	Clinical controls
	.66***	.58***		YSR	PAI-A BOR	YES	Concurrent	
	.55***	.61***		YSR	BPFSC	YES	Concurrent	
	-.04	.28***		YSR	BPFSP	NO	Concurrent	
	.27***	.26***		CBCL	CI-BPD	NO	Concurrent	
	.19* (*d*=.38)	.24** (*d*=.50)	.26** (*d*=.53)	CBCL	CI-BPD (dich.)	NO	Concurrent	Clinical controls
	.37***	.21*		CBCL	PAI-A BOR	NO	Concurrent	
	.24***	.27***		CBCL	BPFSC	NO	Concurrent	
	.30***	.69***		CBCL	BPFSP	YES	Concurrent	
Westen et al. (2003)[a]	.31***	.45***		CBCL	DSM criteria	YES	Concurrent	
	.26***	.50***		CBCL	SWAP-200 A	YES	Concurrent	
Zelkowitz et al. (2007)	.32* (*d*=.68)	.25 (*d*=.51)		CBCL	CDIB-R	NO	(PD assessment preceded ASEBA administration)	Clinical controls

Notes. *$p < .05$ **$p < .01$ *** $p < .001$

a. These studies are reviewed in the section: "All PDs", p. 57.
BPD: Borderline Personality Disorder
YSR: Youth Self-Report (Achenbach, 1991)
SCID-II: Structured Clinical Interview for DSM-III Axis II Personality Disorders (First, Gibbon, Spitzer, Williams, & Benjamin, 1997)
Dich.: dichotomised
CBCL: Child Behaviour Checklist (Achenbach, 1991)
SIDP-IV: Structured Clinical Interview for DSM-IV Personality Disorders (Pfohl, Blum, & Zimmerman, 1995)
TRF: Teacher Report Form (Achenbach, 1991)
BPFSC: Borderline Personality Features Scale for Children (Crick, Murray-Close, & Woods 2005)
BPFSP: Borderline Personality Features Scale for Parents (Sharp, Mosko, Chang, & Ha, 2011)
CI-BPD: Childhood Interview for DSM-IV Borderline Personality Disorder (Zanarini, 2003)
PAI-A BOR: Personality Assessment Inventory for Adolescents (Morey, 2007)
SWAP-200-A: Shedler-Westen Assessment Procedure-200 for Adolescents (Westen & Shedler, 1999a; 1999b)
CDIB-R: Retrospective Diagnostic Interview for Borderlines (Greenman, Gunderson, Cane, & Saltzman, 1986)

As Tables 4 and 5 show, there was a correlation between BPD and most scale scores, but effect sizes varied. The direction of the effect was consistently positive. Overall, the Anxious/Depressed scale had higher concurrent validity compared to the Withdrawn/Depressed scale. Evidence of strong concurrent validity was also reported with regard to the Aggressive and Delinquent Behaviour scales, and all three broadband scales, which consistently differentiated groups with BPD from other clinical non-BPD groups.

Narcissistic PD (NPD)

Three cross-sectional studies reported associations between ASEBA scales and continuous, self-report measures of narcissistic features.

In a study reviewed previously, Penney et al. (2008) used the YSR Aggressive and Delinquent Behaviour scales to validate the MACI Psychopathy scale, the MACI Narcissism or Egotistic Personality scale. This scale is composed of 39 items that can be organised into three factors, namely confidence, exhibitionism/superiority, and conceit/assuredness.

Factor 1 (confidence) was unrelated to Delinquent or Aggressive Behaviour, but factor 2 (exhibitionism/ superiority) was significantly related to both YSR scales, and so was factor 3 (conceit/assuredness), although the correlations with factor 3 were inverse. As mentioned previously, the relatively small sample size, the reliance on self-report measures and the cross-sectional design limit the conclusions that can be drawn from this study.

In another study reviewed before (Underwood et al., 2011), a community cohort of children completed the NPIC (Barry, Frick, & Killian, 2003), a modified version of the adult Narcissistic Personality Inventory (NPI; Raskin & Hall, 1979). Their eighth-grade teachers completed the TRF, from which the Emotional Instability subscale (described previously)

was derived, as proposed by Carlson et al. (2009). Results showed that the correlations between the NPIC and the TRF scales were not significant, as was the case for BPD traits. However, the reliance on self-reports for measuring narcissistic personality features is a limitation. On the other hand, the use of a quite large, typically developing sample, and the inclusion of teacher reports are obvious strengths of this study.

In another community study of young adolescents, Washburn, McMahon, King, Reinecke, and Silver (2004) used the NPI (Raskin & Hall, 1979) to examine the association of narcissistic features with (among other factors) the Anxious/Depressed scale of the YSR. A factor analysis of the NPI resulted in three factors, namely adaptive narcissism, exploitativeness, and exhibitionism. Out of these, only the exhibitionistic factor correlated positively with the anxiety/depression scale, and this relationship remained significant, after accounting for the contributions of other variables (such as gender and self-esteem).

The relevance of this study's findings to our research questions is limited to the Anxious/Depressed scale of the YSR. The size of the effect concerning the association between the Anxious/Depressed scale and the total NPI score was not reported. Therefore, no comparison with the findings of the previous study could be made. In addition, the reliance on self-reports is problematic, and the generalisability of the findings is limited by sampling issues (primarily inner-city young African-American adolescents).

Cluster B and Cluster C PDs

In another publication of a study reviewed previously, Burnette, South, and Reppucci (2007) examined the underlying structure of Cluster B pathology and its association with the YSR. They factor-analysed the borderline, narcissistic, and histrionic symptoms of the SIDP-IV (Pfohl et al., 1997) and obtained a three-factor solution consisting of the *dramatic,* the *vulnerable,* and the *erratic* personality style.

Results showed that all derived factors correlated significantly with the externalising scales, while the correlations with the internalising scales were less strong. However, the *vulnerable* factor, which primarily consisted of BPD characteristics (feelings of emptiness, suicidality, identity disturbance) correlated highly with all externalising and internalising scales, and especially the Anxious/Depressed scale.

Apart from the limitations reported previously in relation to this study, another caveat is that the original Cluster B diagnoses were not used, so the validity estimate reported here was not related to a specific PD type.

The following longitudinal studies looked at Axis II psychopathology in relation to the CBCL-Dysregulation Profile (CBCL-DP), which was initially used to investigate juvenile bipolar disorder (Biederman et al., 1995). This profile is increasingly understood as an indicator of problem severity and overall psychopathology, rather than a predictor of any one specific disorder (Meyer et al., 2009). The CBCL-DP is characterised by co-occurring high scores (namely *T* scores ≥ 70) on the Attention Problems, Aggressive Behaviour, and Anxious/Depressed scales (Althoff, Rettew, Faraone, Boomsma, & Hudziak, 2006; Hudziak, Althoff, Derks, Faraone, & Boomsma, 2005). As a result, it describes children with severe dysregulation and significant elevations in both internalising and externalising difficulties, which can predispose them to complex psychopathology in adulthood.

The first study by Meyer et al. (2009) is a 23-year prospective study of high-risk children (whose mothers had mood disorders) and a control group. As part of the young adult follow-up, participants were assessed for Cluster B PD with the IPDE (Loranger, 1995; Loranger et al., 1997). The study demonstrated that participants meeting the CBCL-DP phenotype (using *T* scores ≥ 60 rather than 70), at least once during childhood and/or adolescence manifested elevated rates of a number of disorders in adulthood, including Cluster B PDs. Specifically, 43% of the adults with Cluster B PDs had the CBCL-PD

phenotype as children whereas only 8% of the children without the phenotype were diagnosed with PD in adulthood. Additionally, there was evidence of construct validity; children with the CBCL-PD phenotype were at greater risk for multiple comorbidity, suicidality and social and occupational impairment, which are also characteristic features of PD. When analyses were repeated using an alternative definition of CBCL-DP, (i.e. a sum of the three scales in question being ≥ 180, as proposed by Faraone, Althoff, Hudziak, Monuteaux, & Biederman, 2005), the findings were similar but the predictive power of the phenotype was decreased somewhat.

Limitations acknowledged in this study include the possible reporting bias of mothers with mood disorders, the application of a lower threshold to establish the CBCL-DP profile (*T* scores ≥ 60 rather than 70) and the small sample size, with only 16 children meeting these less stringent CBCL-DP criteria. Another issue concerns the external validity of the study, as there was an over-representation of mothers with affective illness in the follow-up subsample. On the other hand, the internal validity may have been compromised by the "super-normal" control group employed (children of parents with no past or present psychopathology), which may have resulted in inflated differences between the children with and without CBCL-DP.

The study's prospective design is a clear strength, as is the relatively low attrition rate from the original sample (16.5%). Furthermore, the assessment of participants at regular intervals and the use of an established measure of PD are also evidence of the methodological robustness of this study. At the same time, few children met criteria for repeated CBCL-DP, and the moderate correlations between the CBCL scales over time suggest partial stability of the behaviours captured (Meyer et al., 2009). The higher CBCL-DP rates observed in adolescence can perhaps be attributed to the vulnerabilities associated with this age group.

Halperin et al. (2011) were the second group to examine the role of the CBCL-DP in predicting Axis II pathology. Specifically, they investigated the outcomes of a clinically-referred sample of children with ADHD by assessing for Axis I and Axis II disorders using the SCID-II (First et al., 1997) nine years later. In this study, the CBCL-DP phenotype was used both categorically (*T* scores ≥ 70) and continuously (i.e. by summing the Attention Problems, Aggressive Behaviour and Anxious/Depressed *T* scores).

Results showed that the CBCL-DP phenotype was significantly predictive of Cluster C PD at follow-up, and there was a similar (but not statistically significant) trend in relation to Cluster B rates. The CBCL-DP was also associated with the number of PD diagnoses at follow-up. There was no association with Cluster A PDs. The predictive utility of the CBCL-DP phenotype was assessed further for the presence of Cluster C PDs and "any PD" by examining the CBCL-DP's sensitivity, specificity and predictive power. While sensitivity was generally modest, it is notable that two thirds of those with the phenotype in childhood developed at least one PD in adulthood.

Somewhat similar to Meyer et al.'s (2009) findings, the stability of the categorically-defined phenotype over time was modest, and the correlation across the two assessment periods was quite small. Unlike Meyer's study, the CBCL-DP scores derived from parent ratings in childhood were significantly higher than those obtained via parent CBCL and YSR reports at follow-up. This may reflect the impact of sampling and respondent issues, and could also suggest different psychopathology trajectories for the groups of children studied (children of mothers with mood disorders vs. children with ADHD).

Limitations acknowledged in this study include the substantial (41%) follow-up attrition, the low base rate of CBCL-DP in the sample and the specific nature of the sample (children with ADHD), which limits the external validity of the findings. On the other hand, the study's prospective design and the use of an established PD criterion are significant strengths.

In order to compare this study's results with the results reported by Meyer et al. (2009), the diagnostic efficiency of CBCL-DP was estimated for Cluster B PDs. Table 6 presents the relevant indices.

Table 6. *Predictive utility of the CBCL-Dysregulation Profile for diagnosing later Cluster B Personality Disorders*

Study	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value	Odds ratio	95% CI
Meyer et al. (2009)	.90	.50	.43	.92	9.13	2.2-37.6
Halperin et al. (2011)	.44	.75	.43	.76	2.35	.9-6.1

As Table 6 indicates, the model developed in Meyer et al.'s study has higher sensitivity (erring toward false positives) than specificity (erring toward false negatives), whereas the opposite is true for Halperin et al.'s model. The varying predictive power reported may have resulted from the assessment method for PD (a self-report measure in Meyer et al.'s study vs. a diagnostic interview in Halperin et al.'s study). It may have also resulted from the lower cutoff scores used by Meyer et al. (*T* scores ≥ 60 rather than 70).

In summary, it appears that the CBCL-DP may be a clinically meaningful indicator, describing a relatively stable pattern of difficulty in regulating emotion, behaviour and cognition. This profile was moderately predictive of later PD but can be a useful marker of those at increased risk (Halperin et al., 2011); it is therefore worthy of further investigation. Whether the use of persistent CBCL-DP (namely, meeting the CBCL-DP criteria over multiple time points) will improve the diagnostic efficiency of the phenotype is a question for future research.

In a retrospective study by Goethals, Willigenburg, Buitelaar, and Van Marle (2008) the usefulness of the CBCL was examined in an adult offender patient group. The study's

participants fell into four groups: psychotic offenders without a PD, psychotic offenders with a PD, non-psychotic offenders with a PD, and non-offender psychotic patients without a PD. In this study, the CBCL was scored on the basis of actuarial data in patients' records.

A significant difference was found between psychotic offenders with a Cluster B PD (such as ASPD, NPD) and the non-offender patients with psychosis but no PD; the former group had higher scores on the Delinquent Behaviour and Attention Problems scales. No significant differences were found between the groups on the remaining narrowband scales. Next, a hierarchic cluster analysis was conducted to investigate whether relatively homogeneous patient groups could be formed on the basis of the CBCL score. It was found that all offenders with a PD had significantly higher scores on Externalising and Attention Problems in youth, but for internalising behaviour there was no difference between the groups.

The main methodological limitation of this study is that it was retrospective and based on actuarial data from case note material. Furthermore, it is unclear whether PD diagnoses derived from a structured interview. The small size and unusual characteristics of the sample also limit the conclusions that can be drawn.

All PDs

Two articles presented a study (Dutra, Campbell & Westen, 2004; Westen, Shedler, Durrett, Glass, & Martens, 2003) in which researchers asked approximately 300 experienced psychiatrists and psychologists to describe a randomly selected 14-18 year-old patient in treatment for personality pathology. Clinicians completed the CBCL (parent-report version), and also assessed Axis II pathology. Only 28.4% of patient descriptions met criteria for Axis II disorder, with the remainder describing sub-threshold personality patterns.

Medium or large effect sizes ($r \geq .5$, Cohen, 1992) were found (Dutra et al., 2004) for the correlations between PPD and Total Problems, SZPD and the Withdrawn/Depressed scale, STPD and the Withdrawn/Depressed, Social and Thought Problems scales, in agreement with the findings of Wickline et al. (2012) reported previously. Correlations equal to or larger than .5 were found between ASPD and NPD and Delinquent, Aggressive Behaviour and Externalising and Total Problems, BPD and Total Problems, NPD and Delinquent, Aggressive Behaviour and Externalising Problems, AvPD and Withdrawn/Depressed, Anxious/Depressed and Internalising Problems, and Dependent PD and Anxious/Depressed and Internalising Problems. All PDs correlated significantly with Social Problems and Total Problems. High scores on both of these scales may be a useful global indicator of PD risk, worthy of further empirical investigation.

The 200-item Shedler-Westen Assessment Procedure for Adolescents (SWAP-200-A; Westen & Shedler, 1999a, 1999b) was also used in this project. The SWAP-200-A is a Q-sort instrument designed for use by skilled clinicians to derive dimensions of personality pathology which largely correspond to those of DSM-IV. Westen et al. (2003) reported this instrument's associations with the CBCL Withdrawn/Depressed, Anxious/Depressed, Delinquent and Aggressive Behaviour, Internalising and Externalising Problems scales. It was found that the concurrent validity of these scales was similar; also, their associations with the SWAP-200-A PD scores and the number of Axis II diagnostic criteria met were similar in magnitude. These associations largely replicated the correlations found between the CBCL scales and the PD types reported above.

A significant limitation of the above studies is the reliance on the treating clinician to obtain all information per case, which is likely to have inflated the associations observed due to shared method variance. On the other hand, clinicians are experienced observers, and alongside self-and parent-reports, the use of clinician reports may be a useful additional source of potentially more reliable, elaborated and systematic information (Clark, 2007; Dutra et al., 2004). At the same time, concerns have been expressed

regarding the robustness of the factor structure of the SWAP-200 (Clark, 2007); additional limitations of this study include the low response rate and the high rates of PD comorbidity in this sample, which however resembles adult data (Westen et al., 2003).

Discussion

Summary of findings

The purpose of this review was to evaluate the usefulness of the ASEBA for measuring juvenile PD traits. Most articles identified by the systematic review discussed ASPD/psychopathy and BPD; the scarce evidence regarding the remaining PD types does not allow any definite conclusions to be drawn. In relation to ASPD, consistent evidence of concurrent validity was found for the Aggressive and Delinquent Behaviour and Externalising Problems scales. This is in agreement with previous research, in that Conduct Disorder expressed in the form of aggressive and delinquent behaviours precede ASPD (Robins, 1996). However, the predictive validity of these scales in longitudinal studies was modest. When ASEBA was administered at two or more time points (to operationalise persistent externalising problems), predictive validity increased. With regard to internalising problems, conflicting results were reported, with some studies finding a positive correlation, some a negative correlation and others reporting no significant effect. Furthermore, a moderating effect was detected by some studies, in that the aforementioned associations were not always applicable to females.

In relation to BPD, a wide range of cross-sectional and longitudinal studies consistently reported that BPD traits or diagnoses were associated with higher internalising and externalising problems. In fact, a few studies reported that the association with externalising problems was stronger. These effects were also noticed when individuals with BPD were compared to other clinical groups, suggesting that a combination of particularly high internalising and externalising problems on the ASEBA may be indicative of borderline pathology.

The mixture of internalising and externalising symptoms is in keeping with the clinical picture of BPD, both in youth (Crowell, Beauchaine, & Linehan, 2009; Kernberg et al., 2000) and adulthood (Zanarini et al., 1989). This pattern was detected in both cross-sectional and longitudinal studies. The longitudinal associations in particular add weight to the consistency of the relationships found between Axis I disorders in childhood and adolescence, and the development of PD in adulthood (Cohen, Crawford, Johnson, & Kasen, 2005). However, the effect sizes reported were widely varying, depending on specific design characteristics that are discussed below.

The use of selected ASEBA scales to operationalise dysregulation was one of the most interesting findings. Self-regulation is considered a fundamental developmental process (Althoff et al., 2012), and severe dysregulation of affect, behaviour, and cognition in childhood can set the stage for severe psychopathology in adulthood (Ayer et al., 2009), such as PD and PD-related symptoms, for instance suicidality (Althoff et al., 2006; Volk & Todd, 2007). The CBCL-DP showed evidence of predictive validity in the longitudinal studies reviewed here, but its predictive power appeared dependent on sampling and measurement issues, as well as on the cutoff used to define the CBCL-DP.

Only Westen et al.'s study included Cluster C PDs in their findings, and they reported positive associations with internalising scales. However, as clinicians were the only informants providing information for both constructs, these results are insufficient and it remains uncertain whether individuals with Cluster C PDs would have a different ASEBA profile from individuals with anxiety or other Axis I disorders.

Methodological issues

There are a number of factors that must be taken into account when evaluating the reviewed literature. Studies varied greatly in their methodological characteristics, such as operationalisation of PD, the respondent types providing ASEBA and PD data, the

demographic characteristics and clinical status of the samples. These factors limit comparison between the findings, and are discussed in more detail below.

Sampling and research setting issues

Sample types and sizes varied greatly among the reviewed articles. Few studies included children, and late adolescence was the developmental period studied the most. The majority of studies were conducted in the USA with predominantly White participants. Differences in the ethnic mix of the samples may have accounted for some of the discrepant findings. In addition, the specific clinical nature of the samples in which the use of ASEBA was examined may have had an impact on findings. Studies that included a PD group, a non-PD clinical control and a healthy control group provided stronger evidence about the ability of the ASEBA to discriminate between PD and other clinical problems. The inclusion of a control group is particularly important, given the significant comorbidity of PD with other disorders (Zanarini, Barison, Frankenburg, Reich, & Hudson, 2009).

Furthermore, some studies testing quite complex mediation and moderation models used relatively small sample sizes, whereas other studies were overpowered, and as a result reported associations that were statistically significant but not clinically meaningful. Furthermore, effects related to a study's setting, such as legal involvement during incarceration or the effects of inpatient or camp environment need to be taken into account, as individuals' functioning is apt to vary considerably from one context to another (McConaughy, 1993).

PD measurement issues

Studies that used structured diagnostic interviews to assess PD, and especially those that reported adequate inter-rater reliability tended to provide evidence of a higher quality. However, dimensional measures of PD were also of value. This approach was at times employed in community studies to detect subclinical levels of PD or those 'at risk',

and to compensate for low frequency of clinical cases in the samples, hence maximising statistical power (Paris, 2000). However, who (or what) can provide the most reliable and valid information for assessing PD remains an important open question (Clark, 2007), and research has shown that inter-rater agreement in the PD field is typically modest and variable (Klonsky, Oltmanns, & Turkheimer, 2002; Walters, Moran, Choudhury, Lee, & Mann, 2004). More generally, dissatisfaction with existing assessment practices has been expressed in relation to PD, and PD assessment has been considered "currently inaccurate, largely unreliable, frequently wrong and in need of improvement" (Tyrer et al., 2006, p.51). Furthermore, studies in which diagnoses were based on clinical observation alone have limitations, given the biases inherent in clinical judgment (Dutra et al., 2004).

Use of ASEBA

Studies also varied in their use of ASEBA scales. Whilst some studies administered all items, others administered only a selection of scales. It remains possible that respondents' ratings were affected by the failure to administer the remaining scales (Kosson et al., 2002). Furthermore, to avoid spurious associations due to overlapping item content, ASEBA items were removed on a few occasions. This probably resulted in smaller effect sizes.

ASEBA forms and respondent types

The majority of studies used the CBCL, as caretaker reports tend to be more available in research with young participants. Furthermore, the primary caretaker is usually the person bringing the young person into the mental health service system and is a critical source of information concerning emotional and behavioural problems. However, parental reports may be biased by parents' own psychopathology.

In two studies, the CBCL was completed by the treating clinician. Despite the potential advantages related to clinicians being professionally trained to assess psychopathology, more studies are needed to explore the validity of the CBCL for use by clinicians. Furthermore, in two other studies, the TRF was completed by camp counsellors; again, it can be argued that camp counsellors may not know children well enough to complete the TRF.

Data collected with the YSR may also be biased in that responses to questions about one's own problems may be affected by selective recall, mood, insight deficits, adolescents' willingness to self-disclose their feelings and difficulties, and their desire to present themselves in socially desirable ways (Barnow et al., 2005; DiLallo, Jones, & Westen, 2009).

Consequently, no single procedure or source of information can provide comprehensive assessment of children's behavioural and emotional problems, and there can be no objective measures of such problems independent of human judgment (McConaughy, Achenbach, & Gent, 1988). It is perhaps unsurprising that convergence has been found to be modest between self-report and non-self-report-based assessment of children's emotional and behavioural problems (Kolko, & Kazdin, 1993). As a result, the multi-informant assessment strategy advocated by Achenbach and colleagues (Achenbach et al., 1987) and the inclusion of observational measures whenever possible is expected to increase the reliability and validity of assessment, as well as address the issue of shared method variance.

Shared method variance

The effect of shared method variance was taken into account in a limited number of studies, and was clearly demonstrated in the study by Sharp, Mosko, et al. (2011), where method invariance appeared to have artificially inflated the PD-ASEBA associations

observed. This limitation applies to several studies reviewed here, as they relied exclusively on a single informant to obtain the data from which both the predictor and criterion variables were derived. In this context, evidence deriving from studies using multiple informants is of higher quality.

Design (longitudinal vs. cross-sectional)

The majority of studies presented concurrent associations between personality pathology and ASEBA, which overall tended to be larger compared to longitudinal associations. In cross-sectional studies, ASEBA's 'predictive' utility relates only to statistical prediction and not to prediction over time (Penney et al., 2008). Retrospective reports are also problematic, as they are subject to recall bias and may reflect current, instead of past difficulties (Wright, 2009). To determine the true relationship between childhood presentation and adult personality pathology, prospective longitudinal studies are needed (Paris, 2000).

However, a few caveats need to be noted in relation to these studies. On the one hand, prospective childhood cohort studies can only ever include small numbers of individuals of interest, namely those with or at risk of developing PD (Goethals et al., 2008). As a result, the cost of obtaining a large sample in a community setting to demonstrate potentially small effects can be high and not always an appropriate use of resources. On the other hand, prospective studies using clinical samples follow-up individuals who are typically offered treatment; this is likely to help prevent a downward trajectory from PD traits to subsequent full-blown PD in at least some young people. In any case, multiple assessment points across development are expected to provide stronger evidence for the predictive validity of ASEBA, as was demonstrated in a few studies of persistent problem behaviours.

Data analysis

Although many studies reported statistically significant associations between PD diagnoses and ASEBA scales, only the two studies looking at the CBCL-DP applied rigorous statistical methods to examine the predictive ability and the clinical usefulness of ASEBA for assessing later PD. Furthermore, a limited number of studies explored the effect of gender as a moderating factor in the relationship between ASEBA and PD. As mentioned above, there was some indication of gender-specific differences in the reported associations. Youth's clinical characteristics and age could have also played a moderating role in these associations, but none of the reviewed studies investigated this empirically. Furthermore, the use of advanced data analysis methods can lead to some conclusions regarding the direction of a relationship. For instance, Forsman et al.'s study (2010) demonstrated that when structural equation modelling was used, ASEBA did not predict PD despite the significant cross-lagged correlation between ASEBA at Time 1 and PD at Time 2.

Theoretical and clinical implications

Childhood and adolescent mental illness is a key risk factor for later psychiatric problems (Copeland, Shanahan, Costello, & Angold, 2009), and important questions remain about diagnostic prediction from childhood and adolescence to adulthood. Understanding the construct of PD in youth also raises profound theoretical and measurement issues and call for a strong measurement model.

The need to determine the most appropriate cutoff scores on relevant dimensions for various clinical decisions has been emphasised by Trull (2005), and this clearly applies to the use of ASEBA. Empirically developed, non-arbitrary cutoff points on various scales could guide clinical decisions to conduct further assessments or even implement appropriate interventions (Clark, 2007). In relation to PD, no such ASEBA cutoff point has

been identified to date. However, the studies reviewed here do suggest that it may be possible to use the ASEBA to detect ‘at risk’ children; those children’s ASEBA profile indicates significant dysfunction in multiple domains. This has clear implications for clinical practice, as efficient detection and treatment of emerging personality difficulties is likely to reduce impairment in adulthood with enormous psychosocial benefits.

Limitations of the review

When interpreting the findings of the reviewed studies, the following limitations should be noted. First, with the exception of the two studies which examined the CBCL-DP, the remaining studies were not designed to specifically evaluate the usefulness of ASEBA for assessing PD traits. Therefore on most occasions the evidence reported was inadequate, and the data that could be extracted were limited to tables of correlations between each study’s measures. Another limitation is the substantial methodological heterogeneity amongst these studies, in that they used different instruments, respondent types and adopted different research designs. Therefore it was not possible to quantitatively analyse the findings. Moreover, the research reviewed reflected a monocultural perspective to psychopathology, as all studies were carried out in North America or Europe (and the majority in the USA). Additionally, the adaptive functioning and social competence items of ASEBA were not included in the review; almost no studies provided data in relation to these, and the review focused on the problem items and scales instead.

Despite the above limitations, this review has been the first attempt at extracting and synthesising information about the usefulness of one of the most widely used and highly validated instruments for detecting PD traits in youngsters at risk of developing PD in adulthood, one of the most debilitating of all mental health disorders.

Suggestions for future directions

With one exception, the reviewed studies employed a variable-centred approach to the analysis of the findings. Person-centred approaches to data analysis (e.g. Q factor analysis, latent profile analysis and latent class analysis) can complement the more traditional variable-centred approaches by allowing us to capture the heterogeneity in individual patterns of personality and psychopathology (Althoff et al., 2012). As a result, they could be usefully examined in further PD research.

Furthermore, in order to explore the construct validity of ASEBA, future research should explore whether ASEBA scales are associated with developmental processes and outcomes conceptually related to PD, such as interpersonal aggression and self-harm. The identification of structurally reliable PD scales would also facilitate the investigation of gender differences or even differences across cultures with respect to the manifestation of PD personality traits (Penney et al., 2008).

Moreover, additional large scale prospective studies with frequent data collection points are needed to assess changes over time in the ways in which adolescents' personality and ASEBA problem behaviour scores are related. Future studies should draw participants from a wider range of clinical and community contexts, preferably using nationally representative samples, or oversampling extremes (e.g. high-risk children). Studies investigating typically developing children in particular allow us to examine the developmental precursors of PD and identify predictors of related adjustment problems before they emerge (Underwood et al., 2011). Such research will allow the exploration of the mechanisms by which trajectories of childhood internalising and externalising symptoms predict personality features later in life.

Given the growing trend to view mental illness, especially personality, dimensionally rather than from a categorical perspective (Sharp et al., 2012), there has been increased interest in the dimensional conceptualisation of PDs. Findings on childhood antecedents of adult PDs suggest that childhood and adolescent temperament

and personality traits are significant developmental antecedents for PDs in adulthood (Krueger & Tackett, 2003; Warner et al., 2004). In line with these theoretical and empirical developments and the proposed changes in DSM-V (APA, 2012) future studies could examine the extent to which various maladaptive personality traits (e.g. negative emotionality, introversion, antagonism, disinhibition) are linked to ASEBA scales and items.

A few high-quality studies have been conducted recently in this area (e.g. De Clercq, Van Leeuwen, Van Den Noorgate, De Bolle, & De Fruyt, 2009). These studies used the CBCL together with temperament measures such as the Temperament and Character Inventory (TCI; Cloninger, 1987) and the Dimensional Personality Symptom Item Pool (DIPSI; De Clercq, De Fruyt, & Mervielde, 2003) and reported meaningful associations between CBCL scales and maladaptive personality traits. Whilst some early findings have already been published, they have yet to contribute much to the field.

Finally, this review could be extended to include other widely used instruments, such as the Strengths and Difficulties Questionnaire (SDQ) (Goodman, 1997). In a preliminary search for PD studies using the SDQ, it became evident that compared to the ASEBA, a much more limited number of studies on juvenile PD had used the SDQ. Future research should also evaluate the usefulness of the SDQ to identify young people at risk of developing personality pathology and who may require further assessment.

Conclusions

Although the CBCL has proved imminently useful for tracking Axis I symptoms and behaviour problems, taken together, the findings of this review showed that existing evidence does not provide a compelling case for use of the ASEBA as a screening tool with adequate validity for assessing PD traits and identifying emerging PD. Nonetheless, PD and ASEBA measures were related in meaningful and informative ways, and the ASEBA showed evidence of criterion validity in many studies. It can be concluded that

children and adolescents displaying extremes of functioning across ASEBA syndromes are likely to have significant clinical and psychosocial difficulties and this could be a significant indicator of vulnerability to emerging PD. Consequently, the CBCL could be used in the future as a triage tool—informing practitioners if an in-depth PD assessment would be required.

However, given the dearth of focused studies in this area, more research is needed to investigate ASEBA's psychometric efficiency, and explore the possibility of refining some of its scales to strengthen its predictive power before recommending its use in clinical services.

In any case, single-point-in-time assessment cannot and should not be expected to yield entirely valid PD diagnoses (Clark, 2007), especially so in youth; therefore a more developmentally informed, life-span perspective on PD assessment is required. As Clark (2007) astutely remarks, personality-both adaptive and maladaptive-is too complex to be assessed comprehensively from a single perspective. Therefore, better understanding of PD will require integration of the shared and unique information that can be provided by self-report, carers and teachers, clinicians, observations across settings, extensive case file review, and eventually laboratory data. A comprehensive interdisciplinary, multiple-levels-of-analysis approach holds much promise (Lenzenweger & Cicchetti, 2005), and is also the contextual approach that conforms to the usual standards of clinical assessment (Carr, 2006).

"Learning how information from these various sources can be integrated most validly and usefully likely will challenge researchers for some years to come" (Clark, 2007, p. 236), and is expected to move the PD field forward. With greater understanding of the risk factors, aetiological pathways, and development of PD, researchers and clinicians will be better equipped to develop targeted prevention and intervention programmes, and ultimately lessen the burden and distress caused by these disorders on young people, their significant others and their communities.

References

Achenbach, T. M. (1991). *Integrative Guide for the 1991 CBCL/4-18, YSR, and TRF profiles*. Burlington, VT: University of Vermont, Department of Psychiatry.

Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin, 101*, 213-232.

Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA School-Age Forms and Profiles.* Burlington, VT: University of Vermont, Research Center for Children, Youth, and Families.

Althoff, R. R., Ayer, L. A., Crehan, E. T., Rettew, D. C., Baer, J. R., & Hudziak, J. J. (2012). Temperamental profiles of dysregulated children. *Journal of Child Psychiatry and Human Development, 43*(4), 511-22.

Althoff, R. R., Rettew, D. C., Faraone, S. V., Boomsma, D. I., & Hudziak, J. J. (2006). Latent class analysis shows strong heritability of the child behavior checklist-juvenile bipolar phenotype. *Biological Psychiatry, 60,* 903-911.

American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders* (*4th edition, text revision*). Washington, DC: American Psychiatric Association.

American Psychiatric Association (2012). *Personality and Personality Disorders*. Available: http://www.dsm5.org/PROPOSEDREVISIONS/Pages/PersonalityandPersonalityDisorders.aspx

Andershed, H., Kerr, M., Stattin, H., & Levander, S. (2002). Psychopathic traits in non-referred youths: A new assessment tool. In E. Blauuw & L. Sheridan (Eds.), *Psychopaths: Current International Perspectives* (pp. 131-158). The Hague: Elsevier.

Arens, E. A., Grabe, H. J., Spitzer, C., & Barnow, S. (2011). Testing the biosocial model of borderline personality disorder: Results of a prospective 5-year longitudinal study. *Personality and Mental Health, 5*(1), 29-42.

Ayer, L., Althoff, R., Ivanova, M., Rettew, D., Waxler, E., Sulman, J., & Hudziak, J. (2009). Child Behavior Checklist Juvenile Bipolar Disorder (CBCL-JBD) and CBCL Posttraumatic Stress Problems (CBCL-PTSP) scales are measures of a single dysregulatory syndrome. *Journal of Child Psychology and Psychiatry, 50*, 1291-1300.

Barnow, S., Lucht, M., & Freyberger, H. (2005). Correlates of Aggressive and Delinquent Conduct Problems in Adolescence. *Aggressive Behavior, 31*(1), 24-39.

Barry, C. T., Frick, P. J., & Killian, A. L. (2003). The relation of narcissism and self-esteem to conduct problems in children: A preliminary investigation. *Journal of Clinical Child and Adolescent Psychology, 32,* 139-152.

Bender, D. S., Dolan, R. T., Skodol, A. E., Sanislow, C. A., Dyck, I. R., McGlashan, T. H., et al. (2001). Treatment utilization by patients with personality disorders. *American Journal of Psychiatry, 158,* 295-302.

Biederman, J., Wozniak, J., Kiely, K., Ablon, S., Faraone, S., Mick, E., et al. (1995). CBCL clinical scales discriminate prepubertal children with structured interview-derived diagnosis of mania from those with ADHD. *Journal of the American Academy of Child Adolescent Psychiatry, 34*, 464-471.

Burnette, M. L., & Reppucci, N. D. (2009). Childhood abuse and aggression in girls: The contribution of borderline personality disorder. *Development and Psychopathology, 21*(1), 309-317.

Burnette, M. L., South, S. C., & Reppucci, N. D. (2007). Cluster B personality pathology in incarcerated girls: Structure, comorbidity, and aggression. *Journal of Personality Disorders, 21*(3), 262-272.

Carlson, E. A., Egeland, B., & Sroufe, L. A. (2009). A prospective investigation of the development of borderline personality symptoms. *Development and Psychopathology, 21*(4), 1311-1334.

Carr, A. (2006). *The Handbook of Child and Adolescent Clinical Psychology: A Contextual Approach* (Second Edition). London: Routledge.

Clark, L. A. (2007). Assessment and diagnosis of personality disorder: Perennial issues and an emerging reconceptualization. *Annual Review of Psychology, 58,* 227-57.

Clark, L. A. (2009). Stability and Change in Personality Disorder. *Current Directions in Psychological Science,18*, 27-31.

Cloninger, C. R. (1987). A systematic method for clinical description and classification of personality variants: A Proposal. *Archives of General Psychiatry, 44*, 573-588.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155-159.

Cohen, P. (2008). Child development and personality disorder. *Psychiatric Clinics of North America, 31*(3), 477-493.

Cohen, P., Crawford, T. N., Johnson, J. G., & Kasen, S. (2005). The children in the community study of developmental course of personality disorder. *Journal of Personality Disorder, 9*(5), 466-486.

Cooke, D. J., & Michie, C. (2001). Refining the construct of psychopathy: Towards a hierarchical model. *Psychological Assessment, 13*, 171-188.

Copeland, W. E., Shanahan, L., Costello, J. E., & Angold, A. (2009). Which Childhood and Adolescent Psychiatric Disorders predict which Young Adult Disorders? *Archives of General Psychiatry, 66*(7), 764-772.

Crawford, T. N., Cohen, P., First, M. B., Skodol, A. E., Johnson, J. G., & Kasen, S. (2008). Comorbid Axis I and Axis II disorders in early adolescence. *Archives of General Psychiatry, 65*(6), 641-648.

Crick, N. R., Murray-Close, D., & Woods, K. (2005). Borderline personality features in childhood: A short-term longitudinal study. *Development and Psychopathology, 17*(4), 1051-1070.

Crowell, S. E., Beauchaine, T. P., & Linehan, M. M. (2009). A Biosocial developmental model of borderline personality: An elaboration and extension of Linehan's biosocial theory. *Psychological Bulletin, 135*, 495-510.

De Clercq, B., & De Fruyt, F. (2007). Childhood antecedents of personality disorder. *Current Opinion in Psychiatry, 20*(1), 57-61.

De Clercq, B., De Fruyt, F., & Mervielde, I. (2003). *Construction of the Dimensional Personality Symptom Item pool*. Unpublished manuscript, Ghent University.

De Clercq, B., Van Leeuwen, K., Van Den Noortgate, W., De Bolle, M., & De Fruyt, F. (2009). Childhood personality pathology: Dimensional stability and change. *Development and Psychopathology, 21*(30), 853-869.

DiLallo, J. J., Jones, M., & Westen, D. (2009). Personality subtypes in disruptive adolescent males. *Journal of Nervous and Mental Disease, 197*(1), 15-23.

Dolan, M. C., & Rennie, C. E. (2006). Reliability and validity of the psychopathy checklist: Youth version in a UK sample of conduct disordered boys*. Personality and Individual Differences, 40*(1), 65-75.

Dolan, M. C., & Rennie, C. E. (2007). The relationship between psychopathic traits measured by the Youth Psychopathic trait inventory and psychopathology in a UK sample of conduct disordered boys. *Journal of Adolescence, 30*(4), 601-611.

Dutra, L., Campbell, L., & Westen, D. (2004). Quantifying Clinical Judgment in the Assessment of Adolescent Psychopathology: Reliability, Validity, and Factor Structure of the Child Behavior Checklist for Clinician Report. *Journal of Clinical Psychology, 60*(1), 65-85.

Faraone, S. V., Althoff, R. R., Hudziak, J. J., Monuteaux, M., & Biederman, J. (2005). The CBCL predicts DSM bipolar disorder in children: A receiver operating characteristic curve analysis. *Bipolar Disorder, 7*(6), 518-24.

Ferguson, C. J., Miguel, C. S., & Hartley, R. D. (2009). A multivariate analysis of youth violence and aggression: The influence of family, peers, depression, and media violence. *The Journal of Pediatrics, 155*(6), 904-908.

First, M. B.., Gibbon, M., Spitzer, R. L., Williams, J. B. W., & Benjamin, L. S. (1997). *Structured Clinical Interview for DSM-IV Axis II Personality Disorders, (SCID-II)*. Washington, D.C.: American Psychiatric Press, Inc.

Fite, P. J., Greening, L., & Stoppelbein, L. (2008). Relation between parenting stress and psychopathic traits among children. *Behavioral Sciences & the Law, 26,* 239-248.

Forsman, M., Larsson, H., Andershed, H., & Lichtenstein, P. (2007). The association between persistent disruptive childhood behaviour and the psychopathic personality constellation in adolescence: A twin study. *British Journal of Developmental Psychology,*

25(3), 383-398.

Forsman, M., Lichtenstein, P., Andershed, H., & Larsson, H. (2010). A longitudinal twin study of the direction of effects between psychopathic personality and antisocial behaviour. *Journal of Child Psychology and Psychiatry, 51*(1), 39-47.

Forth, A. E., Kosson, D. S., & Hare R. D. (2004). *The Hare Psychopathy Checklist: Youth Version (PCL-YV) - Rating Guide.* Toronto, Ontario: Multi Health Systems.

Frick, P. J., Bodin, S. D., & Barry, C. T. (2000). Psychopathic traits and conduct problems in community and clinic-referred samples of children: Further development of the Psychopathy Screening Device. *Psychological Assessment, 12*, 382-393.

Geiger, T., & Crick, N. R. (2001). A developmental psychopathology perspective on vulnerability to personality disorders. In R. Ingram & J. M. Price (Eds.), *Vulnerability to Psychopathology: Risk Across the Life Span* (pp.57-102). New York: Guilford.

Goethals, K., Willigenburg, L., Buitelaar, J., & Van Marle, H. (2008). Behaviour problems in childhood and adolescence in psychotic offenders: An exploratory study. *Criminal Behaviour and Mental Health, 18*(3), 153-165.

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A Research Note. *Journal of Child Psychology and Psychiatry, 38,* 581-586.

Greenman, D. A., Gunderson, J. G., Cane, M., & Saltzman, P. R. (1986). An examination of the borderline diagnosis in children. *American Journal of Psychiatry, 143*(8), 998-1003.

Halperin, J. M., Rucklidge, J. J., Powers, R. L., Miller, C. J., & Newcorn, J. H. (2011). Childhood CBCL bipolar profile and adolescent/young adult personality disorders: A 9-year follow-up. *Journal of Affective Disorders, 130*(1-2), 155-161.

Hare, R. D. (1998). The Hare PCL-R: Some issues concerning its use and misuse. *Legal and Criminological Psychology, 3,* 101-122.

Hare, R. D. (2003). *Hare Psychopathy Checklist-Revised (PCL-R), 2nd ed. Technical manual.* North Tonawanda, NY: Multi-Health Systems.

Hudziak, J. J., Althoff, R. R., Derks, E. M., Faraone, S. V., & Boomsma, D. I. (2005). Prevalence and genetic architecture of child behavior checklist-juvenile bipolar disorder. *Biological Psychiatry, 58*, 562-568.

Johnstone, L., & Cooke, D. J. (2004). Psychopathic-like traits in childhood: Conceptual and measurement concerns. *Behavioral Sciences and the Law, 22*, 103-125.

Kernberg, P., Weiner, A., & Bardenstein, K. (2000). *Personality Disorders in Children and Adolescents.* New York: Basic books.

Klonsky, E. D., Oltmanns, T. F., & Turkheimer, E. (2002). Informant-reports of personality disorder: Relation to self-reports and future research directions. *Clinical Psychology: Science and Practice, 9*, 300-11.

Kolko, D. J., & Kazdin, A. E. (1993). Emotional/behavioral problems in clinic and nonclinic children: Correspondence among child, parent, and teacher reports. *Journal of Child Psychology and Psychiatry, 34,* 991-1006.

Kosson, D. S., Cyterski, T. D., Steuerwald, B. L., Neumann, C. S., & Walker-Matthews, S. (2002). The Reliability and Validity or the Psychopathy Checklist: Youth Version (PCL: YV) in Nonincarcerated Adolescent Males. *Psychological Assessment, 14*, 97-109.

Krueger, R. F., & Tackett, J. L. (2003). Personality and psychopathology: Working toward the bigger picture. *Journal of Personality Disorders, 17*, 109-128.

Lenzenweger, M. F., & Cicchetti, D. (2005). Toward a developmental psychopathology approach to borderline personality disorder. *Development and Psychopathology, 17*(4), 893-898.

Lexcen, F. J., Vincent, G. M., & Grisso, T. (2004). Validity and structure of a self-report measure of youth psychopathy. *Behavioral Sciences & the Law, 22,* 69-84.

Loranger, A. W. (1995). *International Personality Disorder Examination (IPDE)*. Geneva: World Health Organization.

Loranger, A. W. (2001). *OMNI Personality Inventories Professional Manual: OMNI Personality Inventory, OMNI-IV Personality Disorder Inventory*. Lutz, FL: Psychological Assessment Resources.

Loranger, A. W., Janca, A., & Sartorius, N. (Eds.) (1997). *Assessment and diagnosis of personality disorders: The International Personality Disorder Examination (IPDE)*. Cambridge: Cambridge University Press.

McConaughy, S. H. (1993). Evaluating behavioral and emotional disorders with the CBCL, TRF, and YSR cross-informant scales. *Journal of Emotional and Behavioral Disorders*, *1*, 40-53.

McConaughy, S. H., Achenbach, T. M., & Gent, C. L. (1988). Multiaxial empirically based assessment: parent, teacher, observational, cognitive, and personality correlates of child behavior profile types for 6- to 11-year-old boys. *Journal of Abnormal Child Psychology, 16,* 485-509.

Meyer, S. E., Carlson, G. A., Youngstrom, E., Ronsaville, D. S., Martinez, P. E., Gold, P. W., et al. (2009). Long-term outcomes of youth who manifested the CBCL-Pediatric Bipolar Disorder phenotype during childhood and/or adolescence. *Journal of Affective Disorders, 113*(3), 227-235.

Miller, A. L., Muehlenkamp, J. J., & Jacobson, C. M. (2008). Fact or fiction: Diagnosing borderline personality disorder in adolescents. *Clinical Psychology Review, 28,* 969-981.

Millon, T. (1993). *Millon Adolescent Clinical Inventory manual.* Minneapolis, MN: National Computer Systems.

Morey, L. C. (2007). *Personality Assessment Inventory-Adolescent: Professional manual.* Odessa, FL: Psychological Assessment Resources.

Murrie, D. C., & Cornell, D. G. (2000). The Millon Adolescent Clinical Inventory and psychopathy. *Journal of Personality Assessment, 75,* 110-125.

National Institute for Mental Health in England (NIMHE) (2003, 23 January). *Personality Disorder: No Longer a Diagnosis of Exclusion*. Retrieved 20 May 2012 from the Department of Health website: http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_4009546

Natsuaki, M. N., Cicchetti, D., & Rogosch, F. A. (2009). Examining the developmental history of child maltreatment, peer relations, and externalizing problems among adolescents with symptoms of paranoid personality disorder*. Development and Psychopathology, 21*(4), 1181-1193.

Paris, J. (2000). Childhood precursors of borderline personality disorder. *Psychiatric Clinics of North America, 23*, 77- 88.

Paternoster, R., & Mazerolle, P. (1994). General strain theory and delinquency: A replication and extension. *Journal of Research in Crime and Delinquency, 31*(3), 235-263.

Penney, S. R., Moretti, M. M., & Da Silva, K. S. (2008). Structural validity of the MACI Psychopathy and Narcissism Scales: Evidence of multidimensionality and implications for use in research and screening. *Journal of Clinical Child and Adolescent Psychology, 37*(2), 422-433.

Pfohl, B., Blum, N., & Zimmerman, M. (1997). *Structured Interview for DSM-IV Personality*. Washington, DC: American Psychiatric Press.

Raskin, R., & Hall, C. S. (1979). A narcissistic personality inventory. *Psychological Reports, 45*, 590.

Robins, L. (1996). Deviant children grown up. *European Child and Adolescent Psychiatry, 5,* 44-46.

Rogosch, F. A. & Cicchetti, D. (2005). Child maltreatment, attention networks, and potential precursors to borderline personality disorder. *Development and Psychopathology, 17(*4), 1071-1089.

Salekin, R. T., Ziegler, T. A., Larrea, M. A., Anthony, V. L., & Bennett, A. D. (2003). Predicting dangerousness with two Millon Adolescent Clinical Inventory psychopathy scales: The importance of egocentric and callous traits. *Journal of Personality Assessment, 80,* 154-163.

Sevecke, K., Lehmkuhl, G., & Krischer, M. K. (2009). Examining relations between psychopathology and psychopathy dimensions among adolescent female and male offenders. *European Child & Adolescent Psychiatry, 18,* 85-95.

Sharp, C., Ha, C., Michonski, J., Venta, A., & Carbone, C. (2012). Borderline personality disorder in adolescents: Evidence in support of the Childhood Interview for DSM-IV Borderline Personality Disorder in a sample of adolescent inpatients. *Comprehensive Psychiatry, 53*(6), 765-774.

Sharp, C., Mosko, O., Chang, B., & Ha, C. (2011). The cross-informant concordance and concurrent validity of the Borderline Personality Disorder Features Scale for Children in a community sample of boys. *Clinical Child Psychology and Psychiatry, 16*(3), 335-349.

Sharp, C., Pane, H., Ha, C., Venta, A., Patel, A. B., Sturek, J., et al. (2011). Theory of mind and emotion regulation difficulties in adolescents with borderline traits. *Journal of the American Academy of Child & Adolescent Psychiatry, 50*(6), 563-573.

Sharp, C., & Romero, C. (2007). Borderline personality disorder: A comparison between children and adults. *The Bulletin of the Menninger Clinic, 71*(2), 121-150.

Shiner, R. L. (2009). The development of personality disorders: Perspectives from normal personality development in childhood and adolescence. *Development and Psychopathology, 21,* 715-734.

Skodol, A. E., Gunderson, J. G., McGlashan, T. H., Dyck, I. R., Stout, R. L., Bender, D. S., et al. (2002). Functional impairment in patients with schizotypal, borderline, avoidant, or obsessive-compulsive personality disorder. *American Journal of Psychiatry, 159*(11), 276-283.

Trull, T. J. (2005). Dimensional models of personality disorder: Coverage and cutoffs. *Journal of Personality Disorder, 19*, 262-82.

Tyrer, P., Coombs, N., Ibrahimi, F., Mathilakath, A., Bajaj, P., Ranger, M., et al. (2006). Critical developments in the assessment of personality disorder. [Special Issue: Assessment, risk and outcome in severe PD]. *British Journal of Psychiatry, 190*(Suppl. 49), 51-59.

Underwood, M. K., Beron, K. J., & Rosen, L. H. (2011). Joint trajectories for social and physical aggression as predictors of adolescent maladjustment: Internalizing symptoms,

rule-breaking behaviors, and borderline and narcissistic personality features. *Development and Psychopathology, 23*(2), 659-678.

Vitacco, M., Neumann, C. S., & Jackson, R. L. (2005). Testing of a four-factor model of psychopathy: Associations with gender, ethnicity, intelligence and violence. *Journal of Consulting and Clinical Psychology, 73*, 466-476.

Volk, H. E., & Todd, R. D. (2007). Does the child behavior checklist juvenile bipolar disorder phenotype identify bipolar disorder? *Biological Psychiatry, 62,* 115-120.

Walters, P., Moran, P., Choudhury, P., Lee, T., & Mann, A. (2004). Screening for personality disorder: A comparison of personality disorder assessment by patients and informants. *International Journal of Methods in Psychiatric Research, 13*, 34-39.

Warner, M. B., Morey, L. C., Finch, J. F., Gunderson, J. G., Skodol, A. E., Sanislow, C. A., et al. (2004). The longitudinal relationship of personality traits and disorders. *Journal of Abnormal Psychology*, *113*, 217-227.

Washburn, J. J., McMahon, S. D., King, C. A., Reinecke, M. A., & Silver, C. (2004). Narcissistic features in young adolescents: Relations to aggression and internalizing symptoms. *Journal of Youth and Adolescence, 33*(3), 247-260.

Westen, D., & Shedler, J. (1999a). Revising and assessing Axis II, part 1: 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 2: Toward an empirically based and clinically useful classification of personality disorders. *American Journal of Psychiatry, 156,* 273-285.

Westen, D., Shedler, J., Durrett, C., Glass, S., & Martens, A. (2003). Personality diagnoses in adolescence: DSM-IV axis II diagnoses and an empirically derived alternative. *American Journal of Psychiatry, 160,* 952-966.

Wickline, V. B., Nowicki, S., Bollini, A. M., & Walker, E. F. (2012). Vocal and facial emotion decoding difficulties relating to social and thought problems: Highlighting schizotypal personality disorder. *Journal of Nonverbal Behavior, 36*(1), 59-77.

Wright, C. (2009). *Where does it all go wrong? A Developmental Perspective on Personality Difficulties*. Unpublished Doctoral Dissertation, University College London, UK.

Zanarini, M. C. (2003). *The Child Interview for DSM-IV Borderline Personality Disorder*. Belmont, MA: McLean Hospital.

Zanarini, M. C., Barison, L. K., Frankenburg, F. R., Reich, D. B., & Hudson, J. I. (2009). Family history study of the familial coaggregation of borderline personality disorder with axis I and nonborderline dramatic cluster axis II disorders. *Journal of Personality Disorders, 23*(4), 357-69.

Zanarini, M. C., Gunderson, J. G., Frankenburg, F. R., & Chauncey, D. L. (1989). The revised diagnostic interview for borderlines: Discriminating BPD from other axis II disorders. *Journal of Personality Disorders, 3*, 10-18.

Zelkowitz, P., Paris, J., Guzder, J., Feldman, R., Roy, C., & Rosval, L. (2007). A five-year follow-up of patients with borderline pathology of childhood. *Journal of Personality Disorders, 21*(6), 664-674.

Part 2: Empirical Paper

Can the Achenbach System of Empirically Based Assessment identify personality disorder traits in young people? A psychometric investigation in a community-based counselling and psychotherapy service for adolescents and young adults

Abstract

Aim: To examine the usefulness of the Achenbach System of Empirically Based Assessment (ASEBA) for detecting personality disorder (PD) traits in young people.

Method: Routine outcome data collected in a community-based psychotherapy service for 1694 young people aged between 12 and 25 years were used in this study. Young people with clinician-rated personality disorder (PD) difficulties were compared to young people with other psychological problems on the Teacher's Report Form (TRF; Achenbach, 1991), the Youth Self-Report (YSR; Achenbach, 1991), the Young Adult Behaviour Checklist (YABCL; Achenbach, 1997) and the Young Adult Self-Report (YASR; Achenbach, 1997). A range of statistical methods were used to compare the two groups. Principal Components Analyses (PCA) and Confirmatory Factor Analyses (CFA) were conducted to explore the PD-related factor structure that potentially underpins the YSR.

Results: Overall, young people with PD problems scored higher on most ASEBA scales, compared to their peers without such problems. The scales that contributed most to group separation were Aggressive Behaviour, Delinquent Behaviour, and Thought Problems. The two YSR-PD scales developed (ASPD and BPD) following the PCA and CFA showed evidence of internal consistency, and of concurrent and convergent validity. However, the model fit indices following the CFA were inadequate.

Conclusions: The ASEBA showed potential for providing useful clinical information about PD-related problems in young people but the findings of this study should be considered preliminary in the absence of a reliable PD criterion.

Introduction

Personality disorders (PDs) are among the most perplexing and debilitating forms of all mental health disorders. In 1980, the Third Edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-III; American Psychiatric Association; APA) introduced a separate Axis II in order to distinguish enduring, trait-like maladaptive personality phenomena from clinical syndromes such as schizophrenia and major depression and to highlight the importance of assessing and treating these disorders (Crawford et al., 2008). PDs are sub-divided into three clusters. Paranoid, schizoid, and schizotypal PDs comprise Cluster A; antisocial, borderline, histrionic, and narcissistic PDs comprise Cluster B; and avoidant, dependent, and obsessive-compulsive PDs comprise Cluster C.

Because of the devastating impact that personality pathology has on individuals and society, PD as a clinical concept has been marked by controversy, which becomes even more pronounced when referring to children and adolescents. On the one hand, objections have been raised concerning the stigmatising effect of PD diagnosis (NIMHE, 2003) as well as the malleability of personality in youth (Seagrave & Grisso, 2002). On the other hand, more recent evidence has demonstrated that personality pathology emerges well before adulthood and that, like adults, some young people's personality difficulties can cause considerable impairment (Shiner, 2009). Therefore, significant benefits may be gained through earlier identification of at-risk young people in terms of prevention and treatment (Crick, Murray-Close, & Woods, 2005; Sharp & Romero, 2007).

A number of juvenile PD assessment tools have been recently developed for this purpose, including clinical interviews, such as the Childhood Interview for DSM-IV Borderline PD (CI-BPD*;* Zanarini, 2003) and parent, clinician and self-report measures, such as the Dimensional Personality Symptom Item Pool (DIPSI; De Clercq, De Fruyt, Van Leuwen, & Mervielde, 2006), and the Shedler-Westen Assessment Procedure-200 for Adolescents (SWAP-200-A; Westen & Shedler, 1999a, 1999b).

However, unlike Axis I disorders, PDs in youth have been largely excluded from standardised assessments (Kernberg, Weiner, & Bardenstein, 2000).

One of the most innovative and informative alternatives to the DSM classification of disorders of childhood and adolescence is the Achenbach's system of empirically based assessment (ASEBA; Achenbach, 2009, www.aseba.org). The ASEBA is one of the most extensively used and best-studied set of instruments for the broadband screening of emotional and behavioural problems, and is also widely used to monitor treatment outcome (Zaslow et al., 2006).

The ASEBA exists in various forms, such as the Child Behaviour Checklist (CBCL; Achenbach, 1991), the Teacher's Report Form (TRF; Achenbach, 1991) and the Youth Self-Report (YSR; Achenbach, 1991). These forms were developed in recognition of the importance of including information from multiple respondents in the assessment of children (Achenbach, McConaughy, & Howell, 1987). The CBCL, TRF, and YSR and their young adult equivalent forms Young Adult Behaviour Checklist (YABCL; Achenbach, 1997) and Young Adult Self-Report (YASR; Achenbach, 1997) are all measures with well-established psychometric properties in clinical, nonclinical, and cross-cultural populations (Achenbach, 1991; Bérubé & Achenbach, 2007; DeGroot, Koot, & Verhulst, 1994).

The problems assessed using the ASEBA can be described on eight syndrome scales and three overall dimensions. The Withdrawn (or Withdrawn/Depressed)[2], Somatic Complaints, and Anxious/Depressed syndromes constitute the Internalising scale, with problems reflecting internal distress. The Delinquent Behaviour[3] and Aggressive Behaviour syndromes constitute the Externalising scale, with problems reflecting conflicts with other people and society's expectations of the individual. The syndrome scales Social

[2] This syndrome scale has been named differently in different forms and versions of the ASEBA. For consistency, we will refer to it here as "Withdrawn".
[3] Referred to as "Rule-breaking behaviour" in the most recent ASEBA versions.

Problems[4], Thought Problems, and Attention Problems were not categorised into a specific group. The broadband Internalising and Externalising Problems are aggregated in a Total Problems score, which has been found to be a significant predictor of referral to mental health services (Ferdinand & Verhulst, 1994; Wiznitzer et al., 1992).

In an effort to link nosologically- and statistically-based taxonomic paradigms, Achenbach and associates (Achenbach, Dumenci, & Rescorla, 2001) used expert opinion to map the ASEBA items to DSM diagnoses in an a priori manner. Using the same pool of items, DSM-oriented scales were created, including Affective Problems, Anxiety Problems, Somatic Complaints, Attention Deficit/Hyperactivity Problems, Oppositional Defiant problems, and Conduct Problems. Although the statistically based and nosologically based paradigms differ, several studies of children and adolescents have reported statistically significant associations between several Axis I diagnoses and related ASEBA scales (Achenbach, Bernstein, & Dumenci, 2005).

In the Adult Self-Report (ASR; Achenbach & Rescorla, 2003) and Adult Behaviour Checklist (ABCL; Achenbach & Rescorla, 2003), which are used for ages 18 to 59, there are two DSM-oriented personality pathology scales, namely Avoidant Personality Problems and Antisocial Personality Problems. It has been also argued that the adult forms contain enough items to also represent much of the schizotypal and obsessive-compulsive PDs (Widiger, 2010). However, there are no Axis II-oriented scales for respondents below 18.

According to recent evidence, specific behavioural and emotional symptoms traditionally described within the internalising–externalising spectrum of psychopathology are significant childhood risk factors for later personality dysfunction (Cohen, 2008; De Clercq & De Fruyt, 2007). In addition, the convergence across different methodological

[4] In the young adult and adult forms, instead of the “Social Problems” scale, there is an “Intrusive” scale.

approaches suggests that internalising and externalising pathology are crucial personality constructs (Westen, Shedler, Bradley, & DeFife, 2012).

Conceptually, several scales and items in the ASEBA forms for children and adolescents relate to personality trait characteristics and could therefore apply to Axis II diagnoses (Kernberg et al., 2000). Illustrative examples are provided by Kernberg et al. (2000), who have estimated that 57% of the questions on the CBCL relate to DSM-IV personality disorder characteristics. In addition, Eggum et al. (2009) used 6 items of the CBCL and TRF Withdrawn scale to operationalise Avoidant PD. However, no independent PD criterion was used to establish the validity of this operationalisation.

Noteworthy exceptions are two longitudinal studies that used the CBCL-Dysregulation Profile (DP). The CBCL-DP is characterised by co-occurring high scores on the Attention Problems, Aggressive Behaviour, and Anxious/Depressed scales (Althoff, Rettew, Faraone, Boomsma, & Hudziak, 2006; Hudziak, Althoff, Derks, Faraone, & Boomsma, 2005). As a result, it describes children with significant elevations in both internalising and externalising problems. The two studies investigating the predictive validity of this phenotype found that children meeting the CBCL-DP criteria were significantly more likely to manifest as adults elevated rates of Cluster B (Meyer et al., 2009) and Cluster C (Halperin et al., 2011) PDs.

Furthermore, Carlson, Egeland and Sroufe (2009) selected from the TRF items representing instability and disturbance in emotional, behavioural, attentional, and relational domains[5]. These are considered to be underlying characteristics of borderline

[5] Emotional items included: “cries a lot”, “nervous high-strung or tense”, “stubborn, sullen, or irritable”, “sudden changes in mood”, “sulks a lot”, “temper tantrums or hot temper”, and “unhappy, sad, or depressed” (α =.77). Behavioural items included: “impulsive, acts without thinking”, “fails to finish things”, “destroys own things”, “accident prone, gets hurt a lot”, “behaves irresponsibly”, explosive and unpredictable behaviour”, and “easily frustrated” (α = .79). Attentional items included: “can’t concentrate, can’t pay attention”, “confused or seems to be in a fog”, “daydreams or gets lost in thoughts”, “absorbed with picking at skin/body”, “sleeps in class”, “stares blankly”, and “inattentive, distracted” (α = .81). Relational items included: “bullying, meanness to others”, “destroys property of others”, “disturbs others”, “doesn’t get along with others”, “gets in fights”, “physically attacks people”, and “threatens people” (α = .87).

personality pathology (Geiger & Crick, 2001). These four domains of functioning in childhood correlated positively with adult borderline symptoms (Carlson et al., 2009). Furthermore, the inter-item reliability of all scales was overall satisfactory. However, these results were not followed up by confirmatory factor analysis, which is recommended for determining the dimensionality of a scale (Cronbach & Shavelson, 2004) and for assessing the measurement model proposed.

As a result, the extent to which the ASEBA can identify PD traits in young people remains an unanswered empirical question. If the ASEBA can identify at-risk young people with relative accuracy, researchers and busy clinicians in community mental health settings could use it as a low cost screening tool to facilitate appropriate use of resources. To date, no studies have thoroughly investigated the psychometric properties of the ASEBA in detecting PD traits in young people.

Research objectives

The Brandon Centre (described below) has been using the ASEBA to systematically collect outcome data using a multi-method measurement strategy as part of an ongoing audit of the mental health services provided. Drawing on the above and acknowledging the need to bridge the gap between academic research and everyday mental health care, the overarching aim of this study is to learn about the nature of Axis II disorders in adolescent populations. Specifically, this study uses the audit data collected at the BC to examine the usefulness of self, clinician and significant-other rated ASEBA scales for detecting PD traits. It seeks to answer the following questions:

a) Do young people with PD traits and symptoms have elevated scores on the ASEBA scales, compared to a non-PD clinical group?

b) Which ASEBA self- and observer-rated syndrome scales are more likely to differentiate young people with personality pathology from those with different mental health problems?

c) How accurately can the TRF instability scales (Carlson et al., 2009) identify young people with PD?

d) How accurately can the equivalent profiles of the CBCL-DP in the YSR, TRF, YASR and YABCL identify young people with PD?

e) Is Kernberg et al.'s (2000) grouping of CBCL items into different PDs empirically supported when using the YSR?

f) Can the YSR items be combined into different clusters which describe types of emerging PD more coherently?

g) Do empirically developed cutoff scores on any of the above subscales have adequate sensitivity and specificity for clinical decision making?

Method

Setting

The Brandon Centre (BC; www.brandon-centre.org.uk) is a well-established, community-based, voluntary sector clinic in Kentish Town, North London, offering a number of services applied to meet the needs of young people aged 12 to 21 years (until recently, the BC has offered services to young people up to the age of 25 years). The services offered include referral and self-referral talking therapies (primarily psychotherapy and counselling), and an advice and information service (mainly on sexual health). As well as parent training for young people with conduct problems (Baruch, Vrouva, & Wells, 2011), the BC also offers Multi-Systemic Therapy (MST; Wells, Adhyaru, Cannon, Lamond, & Baruch, 2010). Previous publications have described in detail the setting and the Centre's approach to treatment (Baruch, 1995; Baruch & Fearon, 2001; Baruch, Fearon, & Gerber, 1998; Baruch, Gerber, & Fearon, 1998).

The BC has been running routine outcome monitoring as part of auditing the service since April 1993 and has been systematically collecting demographic, diagnostic and service use data. Young people who are unwilling or unable to participate are excluded (Baruch, 1995). The perspectives of the person in treatment, a significant other (chosen by the young person), and the clinician are all included. This design is based on a model proposed by Fonagy and Higgitt (1989), which is commonly used in most routine outcome monitoring studies (Baruch & Vrouva, 2010).

Measures

PD Diagnosis

Diagnosis of the young people was based on a slightly modified version of ICD-10 (World Health Organisation, 2010), which includes nine possible diagnoses, all of which are rated on a scale of 0 (None) to 3 (Severe). Following two clinical interviews, therapists with advanced post-graduate clinical training and instructed in the usage of ICD-10 assign one or more relevant diagnoses for each young person and also select a single principal diagnosis (Baruch, 1995). All ICD-10 diagnoses have been organised in an overarching/summary way, rather than identifying the specific conditions within each diagnostic category. As a result, the diagnostic group most relevant to PD is F6, which includes not only PDs, but also disorders of gender identity or sexual orientation, and habit/impulse disorders. The Director of the BC confirmed that the F6 diagnosis was primarily given to young people with PD, but it was also given to a smaller group of young people to describe their gender identity or sexual orientation issues and habit/impulse disorders (G. Baruch, personal communication, June 7, 2012).

The database does not contain data concerning which (if any) specific PD type participants were diagnosed with. However, the PD types that most clinicians considered when assessing young people were Antisocial Personality Disorder (ASPD) and Borderline Personality Disorder (BPD; G. Baruch, personal communication, June 7, 2012). Excluding

a diagnostic grouping for which there were fewer than three positive ratings in total, inter-rater reliability of the diagnoses was reasonably high for the remaining eight groupings, with kappa ranging between 0.6 and 1.0 (Baruch, 1995).

In order to capture the wider range of PD severity, the variable that rated PD as none, mild, moderate, or severe was dichotomised by combining the “mild”, “moderate”, and “severe” classifications into one “PD traits” category, using “none” as the other category. This differentiated between participants with no PD (traits) from those with at least a mild degree of such difficulties, usually below the diagnostic threshold. From this point onward we will refer to this group (with mild, moderate, or severe PD) as the “PD traits” group. As limitations with this approach emerged, in subsequent analysis an alternative operationalisation of PD was used. This is described below (under “Data analysis procedures”).

The following two scales were primarily used for descriptive purposes and are described below.

Severity of Psychosocial Stressors Scale (SPS) for Children and Adolescents

The SPS is taken from Axis IV of the DSM-III-TR (APA, 1987) and involves rating the young person for the severity of psychosocial and environmental stressors on a range of increasing severity from 1 to 6. In several publications of the BC, the median for the population of young people seen at the BC has been reported to be 4 (e.g. Baruch, 1995, Baruch & Fearon, 2001; Baruch et al., 2009). This reflects severe events or circumstances such as divorce of parents, unwanted pregnancy or arrest, or harsh or rejecting parents, chronic life-threatening illness in a parent or multiple foster home placements.

Global Assessment of Functioning Scale (GAF)

The Global Assessment of Functioning Scale (GAF) is a shortened version of the Global Assessment Scale (GAS) and Children's Assessment Scale included in DSM-III TR

(APA, 1987) and DSM-IV (APA, 1994) as Axis V. The therapist assesses the young person's symptomatology and level of functioning according to guidelines on a scale of 1 to 100 of decreasing severity. A score of 70 is normally considered to be the cutoff point between the non-clinical and clinical ranges. GAF scores higher than 70 indicate satisfactory mental health and good overall functioning; scores from 51 to 70 signify mild or moderate impairment or distress; and scores below 51 indicate severe impairment.

Presentation of Problems Form

The therapist also completes a standard form for personal details of the young person (such as demographic, familial, educational characteristics, etc) and the BC's own Presentation of Problems Form comprising 39 items, which can be combined into 25 main problems. The problems describing the young person's current situation are noted as either present or absent (Baruch, 1995). This form was used to create an alternative PD operationalisation (described under "Data analysis procedures").

ASEBA forms

The Youth Self-Report Form (YSR; Achenbach, 1991) was designed for young people between 11 and 18 years old and contains 112 items. The form has been slightly modified by the BC to make it easier to fill out for young people who are not used to American English and also to make it more appropriate for older adolescents; for instance, references to 'kids' were changed to 'young people' (Baruch, 1995). The YSR was completed by the young person at their first or second session (Baruch, 1995).

The Teacher's Report Form (TRF; Achenbach, 1991) is a 113-item form that was developed for obtaining teachers' reports of young people aged between 11 and 18 years. This is because school is a significant context in which young people exhibit normal and problem behaviours and also because teachers, beside parents or other caregivers, are usually the most important adults in young people's lives (Baruch, 1995). For the purposes of

the audit, a significant other chosen by the young person, for instance a peer, a parent, a teacher or a GP completed this form.

The therapist who assesses the young person and provides treatment also fills out the TRF after three appointments. This was introduced nine months after the audit had started because there were concerns about the rate of attrition of significant-others' forms (Baruch, 1995). According to Achenbach (1991) it is crucial that the respondent who completes the TRF has known the young person for at least two months. Clearly this compromises the validity of the therapist completing a form. However, it could be argued that the specialised skill of the clinician in eliciting information as part of history-taking and assessing the young person in the first three sessions gives the clinician a unique insight into the young person's life and difficulties (Baruch & Vrouva, 2010).

The Young Adult Self-Report (YASR; Achenbach, 1997) and the Young Adult Behaviour Checklist (YABCL; Achenbach, 1997) consist of 118 and 115 items respectively, and are the equivalent of the YSR and the TRF for young people aged 18-30 (Achenbach, 1997). These forms have been now replaced by the Adult Self-Report (ASR; Achenbach & Rescorla, 2003) and the Adult Behaviour Checklist (ABCL; Achenbach & Rescorla, 2003) respectively, which incorporate many items of the 1997 editions of the YASR and YABCL. The YASR and YABCL were used at the BC because they were the first post-eighteen forms made available and they were not replaced for consistency purposes (G. Baruch, personal communication, June 21, 2012).

In all ASEBA forms, respondents are instructed to rate the extent to which each item describes the young person now or within the past 6 months (or 2 months in the TRF) by circling a 0 if the item is *not true* of the young person, a 1 if the item is *somewhat or sometimes true*, and a 2 if it is *very true or often true*. At the end of the form, the TRF and the YABCL contain an open-ended item for obtaining reports of additional behavioural/emotional problems not otherwise listed on the forms. Normalised *T* scores provide information about the severity of problems by showing how a young person

compares with normative samples of randomly selected non-referred young people on each scale for each sex/age group, and the Total problems scale yields a *T* score of general psychiatric functioning.

Participants

General characteristics of the BC's service users

About two-thirds of young people seen at the BC are 17 years old and above (Baruch, 1995). This is the target population for the BC's psychotherapy service which is aimed at young people who are too old for CAMHS and for whom NHS adult mental health services are not yet appropriate (Baruch, 1995). The higher percentage of young women having psychotherapy at the BC may partly be accounted for by the existence of the birth control service (Baruch, 1995), although the higher participation of young women is typical of many mental health services. The majority of young people seen at the BC live in the London boroughs of Camden or Islington which are areas of high social and economic deprivation. Nearly 60% either live with a single parent, alone or in a hostel, and over 90% of young people report family problems. Such factors are considered to place young people at greater risk of psychopathology (Baruch, 1995). Indeed, young people usually present with more than one diagnosis and with multiple problems (Baruch, 1995).

Characteristics of data sample

The BC audit database made available for this research's purposes contains demographic and clinical records of young people who were seen at the Centre from 1993 to date. Only data collected at intake were used in this study. Of the 2145 cases, 451 were excluded because of missing data on key variables of interest (e.g. ASEBA scales and/or PD). Of the remaining 1694 young people that were included in the study, YSR item-level data were available for 830 participants and PD and ASEBA scale data were available for

1479 participants. Included and excluded participants did not differ on any variables of interest.

Of the 1694 participants that were included in the study, self-report ASEBA scales (YSR and YASR) were available for 1608 participants, therapist reported scales (TRF and YABCL) were available for 1326 participants and significant-other reported scales (TRF and YABCL) were available for 1125 participants. Multiple informant data were available for 1463 participants.

Data related to the severity of ICD-10 diagnoses were available for 1489 participants. The majority (1062) participants were considered to have no PD traits whereas the remaining 427 participants were considered to have some form of PD-related difficulty, ranging from mild (*n*=201), to moderate (*n*=191), to severe (*n*=35).

Data analysis procedures

Statistical analyses were performed in SPSS Version 19, and MPlus 5.2 (Muthén & Muthén, 2007). Missing data appeared to be randomly scattered throughout groups and predictors. Inspection of the ASEBA distributions revealed that most scales were positively skewed. Square root, logarithm and inverse transformations were attempted but the distributions remained skewed. Therefore, untransformed data were used in all analyses. Given the descriptive and exploratory nature of our analyses, rather than being hypotheses driven, no alpha adjustments were made. Results are presented in two sections: Part I and II.

In Part I, descriptive statistics were used to describe the key characteristics of the sample. Pearson's correlations were used to estimate agreement among respondents, and *t* tests and χ^2 tests were conducted to explore differences between participants with and without PD (traits) on demographic and clinical variables at a bivariate level. Multivariate Analyses of Variance (MANOVAs) were carried out to determine whether young people with or without PD differed on the ASEBA narrowband scales combined linearly.

Because the three broadband dimensions of the ASEBA are summary scores rather than distinct domains of symptomatology, and because the correlations between them were very high, the internalising, externalising and total problems scales were not used in the analyses. All the remaining narrowband scales were used, but analyses were conducted separately for the YSR and the YASR, and the TRF and the YABCL. This is because the YSR and the TRF contain the narrowband scale "Social Problems" which does not exist in the YASR and YABCL, whereas the latter scales contain the "Intrusive" narrowband scale, which in turn does not exist in the YSR and TRF scales. Because sample sizes were unequal, prior probabilities were based on the observed group sizes.

Furthermore, as sample sizes were unequal and the data not normally distributed, the assumption of multivariate normality was untenable and the results of Box's *M* test of equality of covariance matrices could not be trusted. However, because in our data larger samples produced greater variances and covariances, the probability values of MANOVA were likely to be conservative, and therefore significant results could be trusted (Tabachnick & Fidell, 2001). Cook's distance and Leverage values were used to detect influential cases but Cook's distance for all cases fell below the cutoff of 1 (as recommended by Cook & Weisberg, 1982, cited in Field, 2005), and no Leverage value was greater than $3(k+1)/n$, as recommended by Stevens (1992), where $k = 8$ (number of predictors) and n = number of participants.

Pillai's trace *V* was used as it is the most robust statistic to violations of MANOVA's assumptions (Field, 2005). Nevertheless, the significance levels of the MANOVA remained the same regardless of the test statistic used (Pillai's Trace *V*, Wilk's Lambda Λ, Hotelling's Trace *T*, or Roy's Largest Root Θ). For parsimony reasons, only Pillai's *V* is reported in the results.

MANOVAs were followed up with discriminant analyses, all of which revealed one discriminant function that significantly differentiated the two groups (PD vs. non-PD). The

structure matrix, which gives the canonical variate correlation coefficients (Bargmann, 1970, as cited in Field, 2005), was inspected to detect the scales that contributed most to group separation. Loadings less than .5 are not interpreted (Tabachnick & Fidell, 2001).

In order to explore whether the ASEBA contributed to the prediction of PD (traits) status independently of associated demographic and clinical factors, we used hierarchical logistic regression. By entering the variables in a sequential fashion, where statistical overlap existed between blocks of variables, the overlapping variance was apportioned to the prior block. Within each block we chose to use simultaneous regression.

Cronbach's alpha was used to calculate the internal consistency of scales consisting of new sets of ASEBA items. Receiver operating characteristic (ROC) analyses were conducted to investigate the capacity of various scales to differentiate between the PD and non-PD groups, and to identify appropriate thresholds by calculating sensitivity values (proportion of true positives correctly identified) against false positive values (1-specificity, i.e. 1-proportion of true negatives correctly identified) at multiple cutoff scores. Cutoff values that maximised the sum of sensitivity and specificity and minimised overall error were selected.

As questions were raised about the validity of PD diagnosis based on clinician's judgment and in an effort to increase the validity of the critical PD classification, a more robust operationalisation of PD was used in the next part of the results. We were unable to assume with confidence that clinicians using the PD diagnosis in their assessments were reliable in their judgments of its severity, which in the BC assessment protocol was confounded with the category judgment (see "PD diagnosis" section above). Over the historical course of the study, definition of PD changed in both DSM and ICD diagnoses and cutoff points could not be assumed to be reliably applied. In order to overcome this problem a reliable retrospective clinical diagnosis was arrived at using the following algorithm.

The BC collects accurate, comprehensive and reliable check-lists of 25 main presenting problems. The same list was in use throughout the course of the study. Three international PD experts and the director of the BC were asked to review the BC's presentation of problems checklist to identify problems which were in their opinion indicators of individuals with likely PD. Inter-rater agreement was significant but modest, as kappa ranged between .42 and .57. However, substance misuse (alcohol/drugs), and antisocial behaviour (delinquency/conduct problems/violence towards others) were endorsed by three experts; abuse (physical/sexual), and self-harm (suicide attempt, thoughts of deliberate self harm, self-mutilation/self-harm) were endorsed by all four experts. As there was strong agreement concerning the definition of these problems as key indicators of PD, young people presenting with all four problems were considered highly likely to have a PD diagnosis. Therefore a new dichotomous PD variable was created by splitting the sample as follows: young people who had co-occurring substance abuse, antisocial behaviour, self-harm and abuse problems were considered participants with probable emergent PD (n=62), whereas young people who had three or fewer of these problems were operationally defined as not meeting criteria for PD (n=1355). This was considered a conservative diagnostic approach with likely low sensitivity and high specificity. The association of this indicator with that based on clinicians' diagnosis was statistically significant but modest $r(1389) = .19$, $p < .001$.

In Part II, in order to explore the factor structure that potentially underpins the YSR, Principal Components Analyses (PCA) were conducted. The factor solutions were subjected to varimax rotations. We also carried out Confirmatory Factor Analyses (CFA), to test the dimensionality of factors consisting of the surviving items. As the item-level data were categorical, weighted least-squares means and variance adjusted (WLSMV) estimation was used to examine model fit. This method uses weighted least-square parameter estimates from the diagonal of the weight matrix and has been recommended for multivariate non-normal data and for categorical/ordinal variables on the basis of

simulation studies (Muthén, du Toit, & Spisic, 1997). Four indices were invoked to assess the efficacy of these models, including the Comparative Fit Index (CFI; Bentler, 1990), the Tucker-Lewis Index (TLI; Tucker & Lewis, 1973), the Root Mean Square Error of Approximation (RMSEA; Steiger, 1990) and the Weighted Root Mean Square Residual (WRMR; Yu & Muthén, 2002). In order to avoid problems of capitalising on chance fluctuations in the sample, analyses were cross-validated and the sample was randomly split into a model development and a model validation group (Breckler, 1990; Cudeck & Browne, 1983; MacCallum, Roznowski, & Necowitz, 1992).

Sample size estimation

In accordance with Tabachnick and Fidell's (2001) guidelines, we followed a conservative approach and selected the DV (internalising problems) with the smallest demonstrated difference, as reported in the study by Sharp, Ha, Michonski, Venta, & Carbone (2012). In this study, young people meeting criteria for BPD based on the Childhood Interview for DSM-IV Borderline Personality Disorder (CI-BPD; Zanarini, 2003) scored higher on the CBCL internalising scale and the size of this effect (d =.38) fell in the small to medium range (Cohen, 1992). Next, the computer program G*power 3 (Faul, Erdfelder, Lang, & Buchner, 2007) was used to calculate the sample size required to detect this effect size, using standard definitions of alpha (0.05) and power (80%). This resulted in a total sample size of 174. The number of observations contained in the database exceeded this requirement for all informant types. Furthermore, the available sample size was sufficient for the MANOVA and discriminant analyses, as there were more cases than DVs in every cell (Tabachnick & Fidell, 2001). In relation to the hierarchical logistic regression, a power calculation could not be performed as no relevant estimates of effect size were available in the literature. However, the large dataset was highly likely to have a sufficient number of observations for this analysis, too.

Likewise, the available sample size was adequate for the PCA even if the marker variables had low loadings (Tabachnick & Fidell, 2001). In relation to the CFA, it has been suggested (Gagne & Hancock, 2006) that a good basis on which to calculate sample size is the number of indicators per factor (*p/f*) ratio. In accordance with these guidelines, given a *p/f* ratio of 12, a sample of 400 participants is required in order to achieve satisfactory convergence, defined as "requiring no more than 1,100 replications to attain 1,000 fully proper solutions" (Gagne & Hancock, 2006, p.71). Again, the available YSR item-level data exceeded the required sample and thus enables a cross-validation of the derived model in order to test the generalisability of the solution.

Ethics

All data used in this study were collected for routine outcome monitoring. All information provided by young people, their therapists and significant others has been kept confidential and coded to protect anonymity throughout the duration of the audit. Ethical approval was gained from the University College London Research Ethics Committee (see Appendix I).

Results

The results are organised into the following sections. Part I presents a series of bivariate analyses comparing the PD (traits) and non-PD (traits) groups on demographic and clinical characteristics. Then, estimates of the inter-rater agreement between the ASEBA forms are reported. The next section presents a series of MANOVAs and Discriminant Analyses comparing the PD and non-PD (traits) groups on the ASEBA syndrome scales. Logistic regressions are then conducted to identify which ASEBA scales predict PD status independently, after controlling for other known predictors of PD (such as gender and mood disorder). Next, the psychometric properties of the TRF Instability scales (Carlson et al., 2009) and the CBCL-DP are investigated. Part II presents a series

of PCA and CFA that investigate the PD-related factor structure of the YSR. The end of this section examines the reliability and validity of the new YSR scales.

Part I

Demographic and Clinical Comparisons

As shown in Table 1, young people with PD traits (namely with mild, moderate or severe PD) were older, tended to have more problems and diagnoses, showed higher levels of severity of psychosocial stressors, and a lower level of general functioning. Furthermore, there were more females in the PD traits group, and young people in this group were more likely to have a principal diagnosis of neurotic, stress-related or somatoform disorder, and obviously PD, and less likely to have a principal diagnosis of depressive or other mood disorder.

Table 1. *Differences between young people with and without PD traits on demographic and clinical characteristics*

Demographic and clinical characteristics of the participants	No PD traits (*n*=1062)	PD traits (*n*=427)	χ^2 or *t* test
Mean age, years (*SD*) (min. 12- max. 25)	18.1 (3.1)	18.9 (2.8)	*t*(1490)=4.5***
Percentage female	779 (60.5%)	260 (73.4%)	χ^2(1,*N*=1492) =24***
Percentage ethnic minorities	277 (26.9%)	101 (23.8%)	χ^2(1, *N*=1454) =1.5, *ns.*
Percentage principal diagnosis: depressive or other mood disorder (F3)	547 (53.7%)	189 (18.2%)	χ^2(1, *N*=1441) =9.4**
Percentage principal diagnosis: neurotic, stress-related or somatoform disorder (F4, F9.3)	293 (28.8%)	77 (44.8%)	χ^2(1,*N*=1441) =17.3***
Percentage principal diagnosis: adolescent or adult personality disorder (F6)	0	95 (22.5%)	χ^2(1,*N*=1441) =245.6***
Percentage principal diagnosis: hyperkinetic or conduct disorder (F9, F9.2)	81 (7.9%)	22 (5.2%)	χ^2=(1,*N*=1441) =3.4, *ns.*
Percentage principal diagnosis: syndromes with physiological symptoms (F5)	45 (4.4%)	13 (3.1%)	χ^2=(1,*N*=1441) =1.4, *ns.*
Percentage principal diagnosis: substance abuse (F10-F19)	20 (2%)	15 (3.6%)	χ^2 (1, *N*=1441) =3.2, *ns.*
Percentage principal diagnosis psychosis, organic syndrome, pervasive developmental disorder or mental handicap (D0, F2, F7, F8.4)	13 (1.3%)	7 (1.7%)	χ^2=(1,*N*=1441) =.3, *ns.*
Percentage principal diagnosis: specific developmental disorder	10 (1%)	1 (.2%)	χ^2=(1,*N*=1441) =2.2, *ns.*
Percentage principal diagnosis: other disorder with childhood onset (F9.4 – F9.8).	10 (1%)	3 (.7%)	χ^2=(1,*N*=1441) =.2, *ns.*
Mean number of diagnoses (*SD*) (min. 0 - max. 9)	2.6 (1.1)	4.7 (1.6)	*t*(1490)=28.9***
Mean number of problems (*SD*) (min. 1- max. 17)	5.8 (2.9)	7.6 (2.7)	*t*(1463)=11.1***
Mean rating for Severity of Psycho-Social Stressors (SPSS) scale (min. 1- max. 6)	3.6 (1.1)	3.8 (1.0)	*t*(1458)=2.9**
Mean score on Global Assessment of Functioning (GAF) scale (1-100) (*SD*) (actual range 2-85)	56.2 (9.6)	48.4 (11.6)	*t*(1472)=13.3***

Note. Percentages may not add up to total due to missing data.
*p<.05 **p<.01 ***p<.001

Table 2 presents the comparisons between young people with and without PD (using the alternative, expert-defined criteria).

Table 2. *Differences between young people with and without PD on demographic and clinical characteristics (PD was operationalised using the alternative, expert-defined criteria)*

Demographic and clinical characteristics of the participants	No PD (*n*=1355)	PD (*n*=62)	χ^2 or *t* test
Mean age, years (*SD*) (min. 12- max. 25)	18.4 (3.1)	18.6 (2.4)	*t*(1415)=.67***
Percentage female	951 (70.2%)	46 (74.2%)	χ^2(1,*N*=1417) =.46, *ns.*
Percentage ethnic minorities	343 (25.9%)	14 (22.6%)	χ^2(1, *N*=1384) =.35, *ns.*
Percentage principal diagnosis: depressive or other mood disorder (F3)	678 (51.7%)	23 (37.7%)	χ^2(1, *N*=1373) =4.55*
Percentage principal diagnosis: neurotic, stress- related or somatoform disorder (F4, F9.3)	334 (25.5%)	16 (26.2%)	χ^2(1,*N*=1373) =.02, *ns*.
Percentage principal diagnosis: adolescent or adult personality disorder (F6)	75 (5.7)%	10 (16.4%)	χ^2(1,*N*=1373) =11.44***
Percentage principal diagnosis: hyperkinetic or conduct disorder (F9, F9.2)	99 (7.5%)	5 (8.2%)	χ^2=(1,*N*=1373) =.35, *ns.*
Percentage principal diagnosis: syndromes with physiological symptoms (F5)	53 (4%)	3 (4.9%)	χ^2=(1,*N*=1373) =.12, *ns.*
Percentage principal diagnosis: substance abuse (F10-F19)	33 (2.5%)	2 (3.3%)	χ^2 (1, *N*=1373) =.14, *ns.*
Percentage principal diagnosis psychosis, organic syndrome, pervasive developmental disorder or mental handicap (D0, F2, F7, F8.4)	18 (1.4%)	1 (1.6%)	χ^2=(1,*N*=1373) =.03, *ns.*
Percentage principal diagnosis: specific developmental disorder	10 (.8%)	0	χ^2=(1,*N*=1373) =.47, *ns.*
Percentage principal diagnosis: other disorder with childhood onset (F9.4 – F9.8).	12 (.9%)	1(1.6%)	χ^2=(1,*N*=1373) =.33, *ns.*
Mean number of diagnoses (*SD*) (min. 0 - max. 9)	3.1 (1.5)	4.9 (1.9)	*t*(1414)=9.1***
Mean number of problems (*SD*) (min. 1- max. 17)	6.2 (2.8)	11.7 (2.4)	*t*(1415)15.5***
Mean rating for Severity of Psycho-Social Stressors (SPSS) scale (min. 1- max. 6)	3.7 (1.1)	4.4 (.5)	*t*(1396)=5.7***
Mean score on Global Assessment of Functioning (GAF) scale (1-100) (*SD*) (actual range 2-85)	54.4 (9.6)	44.6(10.4)	*t*(1402)=6.9***

Note. Percentages may not add up to total due to missing data.
*p<.05 **p<.01 ***p<.001

Inter-rater agreement

Agreement ratings as assessed by Pearson's correlations were, according to Cohen's standards (Cohen, 1977), moderate to high. Between self and therapist the highest agreement was observed for Externalising Problems [$r(1633) = .49$, $p < .001$], followed by Internalising Problems [$r(1633) = .46$, $p < .001$] and Total Problems [$r\ (1633) = .40$, $p < .001$]. As concerns agreement between self and significant-other ratings, the highest agreement was observed for Externalising Problems [$r(1377) = .51$, $p < .001$], followed by Internalising Problems [$r(1377) = .42$, $p < .001$] and Total Problems [$r(1377) = .38$, $p < .001$]. Agreement between therapist and significant-other ratings was also highest for Externalising Problems [$r(1213) = .47$, $p < .001$], followed by Total Problems [$r\ (1213) = .40$, $p < .001$] and Internalising Problems [$r(1213) = .35$, $p < .001$]. The modest correspondence found between self and other reports is in line with findings of other studies using the ASEBA (e.g. Achenbach et al., 1987; Kolko & Kazdin, 1994; Phares, Compas, & Howell, 1989).

ASEBA comparisons

Tables 3 and 4 present the means and standard deviations (*SD*s) of participants' *T* scores on self-reported ASEBA narrowband and broadband scales.

Table 3. *Descriptive statistics for the ASEBA scales (T scores) at intake according to PD traits status*

	YSR/YASR (self-report) Mean (SD)		TRF/YABCL (therapist) Mean (SD)		TRF/YABCL (significant other) Mean (SD)	
	No PD traits (*n*=962)	PD traits (*n*=408)	No PD traits (*n*=820)	PD traits (*n*=343)	No PD traits (*n*=701)	PD traits (*n*=297)
Withdrawn	61.4 (10.2)	63.8 (9.2)	63.4 (8.1)	64.6 (7.9)	61.5 (9.1)	69.9 (8.2)
Somatic Complaints	61.5 (9.7)	63.1 (10)	59.3 (9.3)	59.9 (9.1)	63.9 (10.7)	64.6 (10.6)
Anxious/Depressed	66.9 (12.3)	70.1 (11.5)	72.4 (9.8)	75.0 (9.5)	68.7 (10.7)	71.4 (10.3)
Social Problems (only YSR)	56.9 (8.9)	59 (7.6)	65.9 (6.7)	68.3 (6.7)	64.2 (8.1)	65.4 (7.8)
Intrusive (only YASR)	55.3 (7.2)	56.9 (6.9)	56.4 (8.2)	59.5 (7.4)	57.5 (8.2)	59.7 (8.1)
Thought Problems	59.8 (10.8)	62.7 (19.6)	64.5 (7.7)	64.5 (7.7)	63.8 (9.9)	67 (9.6)
Attention Problems	62.2 (9.2)	64.6 (9.2)	60.7 (7.6)	63.3 (7.0)	60.0 (8.7)	62.3 (7.6)
Delinquent Behaviour	59.8 (8.7)	62.4 (8.2)	59.3 (8.9)	63.6 (7.6)	60.5 (8.3)	63.1 (7.8)
Aggressive Behaviour	59.6 (9.10)	61.7 (9.3)	61.3 (8.6)	65.6 (6.7)	61.5 (8.0)	63.2 (7.9)
Internalising Problems	65.2 (10.8)	68.2 (10.5)	70.9 (9.1)	73.4 (8.9)	68.2 (10.6)	70.9 (10.3)
Externalising Problems	58.8 (10.2)	62 (9.8)	60.6 (7.3)	64.7 (7.1)	60.9 (8.2)	63.1 (7.5)
Total Problems	63.6 (9.6)	67.2 (9.5)	65.0 (6.5)	68.7 (6.1)	64.7 (8.2)	67.5 (8.1)

YSR: Youth Self-Report; YASR: Young Adult Self-Report; TRF: Teacher's Report Form; YABCL: Young Adult Behaviour Checklist

Table 4. *Descriptive statistics for the ASEBA scales (T scores) according to PD status* (*PD was operationalised using the alternative, expert-defined criteria)*

	YSR/YASR (self-report) Mean (SD)		TRF/YABCL (therapist) Mean (SD)		TRF/YABCL (significant-other) Mean (SD)	
	No PD (*n*=1291)	PD (*n*=57)	No PD (*n*=1047)	PD (*n*=54)	No PD (*n*=905)	PD (*n*=35)
Withdrawn	62.2(9.6)	63.1(9.1)	63.9(8.1)	63.1(7.1)	62.1(8.6)	66.5(8.2)
Somatic Complaints	61.6(10.0)	67.4(9.6)	59.5(10.3)	62(9.2)	64.1(11.2)	67.3(10.7)
Anxious/Depressed	67.8(11.7)	70.8(10.9)	73.2(9.7)	74.3(8.3)	69.5(10.4)	75.3(9.3)
Social Problems (only YSR)	57.5(10.1)	58.9(8.6)	66.5(6.9)	70(6.3)	64.6(8.4)	67.7(7.9)
Intrusive (only YASR)	55.6(7.1)	57.9(6.7)	57.3(7.8)	60.3(7.8)	57.8(7.9)	63.0(10.6)
Thought Problems	60.2(11.5)	69.9(9.8)	65.7(9.6)	70.9(8.4)	64.4(10.5)	74.8(9.6)
Attention Problems	62.7(9.2)	67.8(8.4)	61.3(10.4)	66.4(7.1)	60.5(11.3)	68.3(7.8)
Delinquent Behaviour	60.1(8.4)	69.2(8.2)	60.1(9.7)	69.9(7.8)	61.0(8.0)	65.9(6.4)
Aggressive Behaviour	59.9(9.7)	68.7(9.1)	62.1(9.5)	71.4(7.1)	61.8(9.9)	67.9(7.9)
Internalising Problems	66.1(10.7)	69.5(8.7)	71.7(9.2)	72.7(8.1)	68.9(10.6)	75.3(8.8)
Externalising Problems	59.4(10.1)	69.6(9.6)	61.4(7.3)	70.1(7.1)	61.4(7.9)	67.2(7.4)
Total Problems	64.4(9.6)	72.1(8.9)	65.9(8.1)	72.1(6.3)	65.4(8.5)	72.9(8.2)

YSR: Youth Self-Report; YASR: Young Adult Self-Report; TRF: Teacher's Report Form; YABCL: Young Adult Behaviour Checklist

Tables 3 and 4 show that the PD (traits) group scored higher than the non-PD (traits) group on all scales, with the exclusion of the Withdrawn syndrome (TRF/YABCL therapist report, Table 4), although the difference was inconsequential.

YSR (self-report)

Participants with PD traits (n = 206) scored higher than those without PD traits (n = 556) on the YSR narrowband scales, V = .05, F (8, 753) = 4.52, p <.001. The discriminant analysis following up the MANOVA revealed one discriminant function, which significantly differentiated the two groups, Λ=.95, $\chi^2(8)$=35.42, p < .001, and the canonical R^2 was modest (.05). The scales that contributed most to group separation were Anxious/ Depressed (r = .87), Withdrawn (r = .79), Attention Problems (r = .59), Delinquent Behaviour(r = .55) and Social Problems (r = .52). Using this function, 72.6% of original grouped cases were correctly classified.

Using the alternative expert-defined criteria, participants with PD (*n* = 23) scored higher than those without PD (*n* = 702) on the YSR narrowband scales, *V* = .05, *F* (8, 716) = 4.63, *p* <.001. The discriminant function significantly differentiated the two groups, Λ=.95, χ^2(8)=36.24, *p* < .001, and the canonical R^2 was again modest (.05). The scales that contributed most to group separation were Thought Problems (*r* = .69) and Delinquent Behaviour (*r* = .67). Using this function, 96.6% of original grouped cases were correctly classified.

YASR (self-report)

Participants with PD traits (*n* = 202) scored higher than those without PD traits (*n* = 406) on the YASR narrowband scales, *V* = .04, *F* (8, 659) = 3.33, *p* <.001. The discriminant function significantly differentiated the two groups, Λ=.96, χ^2(8)=26.25, *p*<.001, and the canonical R^2 was again modest (.04). The scales that contributed most to group separation were Delinquent Behaviour (*r* = .84), Thought Problems (*r* = .74), Aggressive Behaviour (*r* = .63), Attention Problems (*r* = .5) and Intrusive (*r* = .5). Using this function, 69.6 % of original grouped cases were correctly classified.

Using the alternative expert-defined criteria, participants with PD (*n* = 34) scored higher than those without PD (*n* = 589) on the YASR narrowband scales, *V* = .05, *F* (8, 614) = 10.04, *p* <.001. The discriminant function significantly differentiated the two groups, Λ=.88, χ^2(8)=75.88, *p* < .001, and the canonical R^2 was .05. The scales that contributed most to group separation were Delinquent Behaviour (*r* = .82), Aggressive Behaviour (*r* = .70) and Thought Problems (*r* = .59). Using this function, 94.4% of original grouped cases were correctly classified.

TRF (therapist report)

Participants with PD traits (*n* = 153) scored higher than those without PD traits (*n* = 404) on the TRF (therapist report) narrowband scales, *V* = .11, *F* (8, 548) = 8.45, *p* <.001. The discriminant function significantly differentiated the two groups, Λ=.89, χ^2(8)=64.43, *p*<.001, and the canonical R^2 was .11. The scales that contributed most to group separation were Delinquent Behaviour (*r* = .72), Thought Problems (*r* = .68) and Aggressive Behaviour (*r* = .53). Using this function, 74.9% of original grouped cases were correctly classified.

Using the alternative expert-defined criteria, participants with PD (*n* = 23) scored higher than those without PD (*n* = 506) on the TRF narrowband scales, *V* = .05, *F* (8, 520) = 7.15, *p* <.001. The discriminant function significantly differentiated the two groups, Λ=.90, χ^2(8)=54.55, *p* < .001, and the canonical R^2 was .05. The scales that contributed most to group separation were Delinquent Behaviour (*r* = .87) and Aggressive Behaviour (*r* = .64). Using this function, 94.7% of original grouped cases were correctly classified.

YABCL (therapist report)

Participants with PD traits (*n* = 190) scored higher than those without PD traits (*n* = 416) on the YABCL (therapist report) narrowband scales, *V* = .16, *F* (8, 597) = 14.09, *p* <.001. The discriminant function significantly differentiated the two groups, Λ=.84, χ^2(8)= 103.8, *p*<.001, and the canonical R^2 was .16. The scales that contributed most to group separation were Aggressive Behaviour (*r* = .79), Thought Problems (*r* = .67) and Delinquent Behaviour (*r* = .57). Using this function, 74.3% of original grouped cases were correctly classified.

Using the alternative expert-defined criteria, participants with PD (*n* = 31) scored higher than those without PD (*n* = 541) on the YABCL narrowband scales, *V* = .13, *F* (8, 563) = 10.91, *p* <.001. The discriminant function significantly differentiated the two groups, Λ=.87, χ^2(8)=81.58, *p* < .001, and the canonical R^2 was .13. The scales that contributed

most to group separation were Aggressive Behaviour (*r* = .84) and Delinquent Behaviour (*r* = .66). Using this function, 94.4% of original grouped cases were correctly classified.

TRF (significant-other report)

Participants with PD traits (*n* = 152) scored higher than those without PD traits (*n* = 361) on the TRF (significant-other report), *V* = .06, *F* (8, 504) = 3.83, *p* <.001. The discriminant function significantly differentiated the two groups, Λ=.94, χ^2(8)= 29.9, *p* <.001, and the canonical R^2 was .06. The scales that contributed most to group separation were Withdrawn (*r* = .65), Delinquent Behaviour (*r* = .63), Thought Problems (*r* = .58), and Anxious/Depressed (*r* = .50). Using this function, 70.8% of original grouped cases were correctly classified.

Using the alternative expert-defined criteria, participants with PD (*n* = 16) scored higher than those without PD (*n* = 472) on the TRF narrowband scales, *V* = .04, *F* (8, 479) = 2.76, *p* <.001. The discriminant function significantly differentiated the two groups, Λ=.96, χ^2(8)=21.8, *p* = .005, and the canonical R^2 was modest (.04). The scales that contributed most to group separation were Thought Problems (*r* = .89) and Attention Problems (*r* = .61). Using this function, 96.7% of original grouped cases were correctly classified.

YABCL (significant-other report)

Participants with PD traits (*n* = 145) scored higher than those without PD traits (*n* = 340) on the TRF on the YABCL (significant-other report), *V* = .05, *F* (8, 476) = 3.07, *p* =.002. The discriminant function significantly differentiated the two groups, Λ=.95, χ^2(8)= 24.1, *p* = .002, and the canonical R^2 was .05. The scales that contributed most to group separation were Thought Problems (*r* = .70), Aggressive Behaviour (*r* = .69), Attention

Problems (r = .67), Aggressive Problems (r = .65) and Intrusive (r = .55). Using this function, 69.3% of original grouped cases were correctly classified.

Using the alternative expert-defined criteria, participants with PD (n = 19) scored higher than those without PD (n = 433) on the YABCL narrowband scales, V = .09, F (8, 479) = 5.37, p <.001. The discriminant function significantly differentiated the two groups, Λ=.91, $\chi^2(8)$=41.25, p < .001, and the canonical R^2 was .09. The scales that contributed most to group separation were Attention Problems (r = .79), Aggressive Behaviour (r = .73), Delinquent Behaviour (r = .68) and Thought Problems (r = .66). Using this function, 95.4% of original grouped cases were correctly classified.

The above results indicate that the PD traits and emergent PD groups had more problems overall, and that the scales that most consistently contributed to group separation were Delinquent Behaviour, Thought Problems and Aggressive Behaviour. Furthermore, when PD was defined using the alternative, expert-defined criteria, prediction was improved.

Group differences in ASEBA after controlling for related factors

Next, hierarchical logistic regressions were conducted to identify the scales that make an independent contribution to the prediction of PD after controlling for other associated factors (see Tables 1 and 2). Age, gender, a principal diagnosis of neurotic, stress-related or somatoform disorder and a principal diagnosis of mood disorder were entered in the first block when predicting PD traits, whereas age and a principal diagnosis of mood disorder were entered in the first block when predicting PD (using expert-defined criteria). In the second block, all ASEBA syndrome scales were used, apart from the Social Problems and Intrusive scales. Because these scales are unique to the YSR/TRF and YASR/YABCL respectively, they were omitted from the regression models for parsimony reasons, in order to combine the adolescent and young adult data into one analysis. The

omission of these scales was considered inconsequential, given their relatively minor contribution to group separation following the discriminant analyses presented above.

Self-reports-PD traits

Using age, gender, neurotic, stress-related or somatoform disorder and mood disorder as predictors, the χ^2 test of goodness of fit was significantly different from zero at the end of the first step, $\chi^2(4)=104.6$, $p<.001$, Cox & Snell $R^2 = .073$, Nagelkerke $R^2=.104$. After step 1, with the seven ASEBA syndromes in the equation (namely Withdrawn, Somatic Complaints, Anxious/Depressed, Thought Problems, Attention Problems, Delinquent Behaviour and Aggressive Behaviour), the model improved reliably, $\chi^2(11) = 145$, $p<.001$, Cox & Snell $R^2 = .10$, Nagelkerke $R^2=.14$.

Table 1 in Appendix II shows the five independent variables that made a significant contribution: age, female gender, neurotic, stress-related or somatoform, mood disorder and Delinquent Behaviour. All these regression coefficients were positive and predicted PD traits. The remaining ASEBA scales did not emerge from this analysis as significant predictors because their effects were explained more powerfully by other variables in the equation.

Self-reports-PD (*using expert-defined criteria*)

Using age and mood disorder as predictors, the χ^2 test of goodness of fit was insignificant, $\chi^2(2) =4.36$, $p= .113$, *ns.* However, with the seven ASEBA syndromes in the equation, the model improved reliably, $\chi^2(9) = 89.6$, $p<.001$, Cox & Snell $R^2 = .06$, Nagelkerke $R^2 = .22$.

Table 2 in Appendix II shows the four independent variables that made a significant contribution: age, Anxious/Depressed, Thought Problems, and Delinquent Behaviour. All

regression coefficients were positive, apart from the Anxious/Depressed syndrome. This means that PD was predicted by lower Anxious/Depressed scores.

Therapist-reports-PD traits

Using age, gender, neurotic, stress-related or somatoform disorder and mood disorder as predictors, the χ^2 test of goodness of fit was significantly different from zero at the end of the first step, χ^2 (4) =85.2, p<.001, Cox & Snell R^2 = .073, Nagelkerke R^2=.104. After step 1, with the seven ASEBA syndromes in the equation, the model improved reliably, χ^2 (11) = 214, p<.001, Cox & Snell R^2 = .17, Nagelkerke R^2=.25.

Table 3 in Appendix II shows the eight independent variables that made a significant contribution: age, female gender, neurotic, stress-related or somatoform, mood disorder, Thought Problems, Somatic Complaints, Delinquent Behaviour, and Aggressive Behaviour. All these regression coefficients were positive, apart from Somatic Complaints, which means that lower, rather than higher scores on this scale predicted PD traits.

Therapist-reports -PD (*using expert-defined criteria*)

Using age and mood disorder as predictors, the χ^2 test of goodness of fit was insignificant, χ^2 (2) =4.26, p= .119, *ns.* However, with the seven ASEBA syndromes in the equation, the model improved reliably, χ^2 (9) = 87.1, p<.001, Cox & Snell R^2 = .08, Nagelkerke R^2 =.24.

Table 4 in Appendix II shows the two independent variables that made a significant contribution: Aggressive Behaviour and Delinquent Behaviour, with positive regression coefficients, which means that high scores on these scales predicted PD.

Significant-other reports-PD traits

Using age, gender, neurotic, stress-related or somatoform disorder and mood disorder as predictors, the χ^2 test of goodness of fit was significantly different from zero at the end of the first step, $\chi^2(4) = 67.4$, $p < .001$, Cox & Snell $R^2 = .067$, Nagelkerke $R^2 = .096$. After step 1, with the seven ASEBA syndromes in the equation, the model improved reliably, $\chi^2(11) = 112.5$, $p < .001$, Cox & Snell $R^2 = .11$, Nagelkerke $R^2 = .156$.

Table 5 in Appendix II shows the seven independent variables that made a significant contribution: age, female gender, neurotic, stress-related or somatoform, mood disorder, Withdrawn, Thought Problems, and Delinquent Behaviour. All regression coefficients were positive and predicted PD traits.

Significant-other reports -PD (*using expert-defined criteria*)

Using age and mood disorder as predictors, the χ^2 test of goodness of fit was insignificant, $\chi^2(2) = 1.5$, $p = .47$, *ns.* However, with the seven ASEBA syndromes in the equation, the model improved reliably, $\chi^2(9) = 44.3$, $p < .001$, Cox & Snell $R^2 = .047$, Nagelkerke $R^2 = .174$.

Table 6 in Appendix II shows the two independent variables that made a significant contribution: Thought Problems and Attention Problems, with positive regression coefficients, which means that high scores predicted PD.

Carlson et al.'s (2009) TRF Instability scales

Next, we investigated the TRF scales representing instability and disturbance in emotional, attentional, behavioural, and relational domains as described by Carlson et al. (2009). The inter-item reliability coefficients (alpha) of these scales ranged, according to Cohen's standards (Cohen, 1977), from questionable ($.7 > \alpha \geq .6$) to good ($.9 > \alpha \geq .8$) and are presented in Table 5. Tables 6 and 7 present the means and SDs of participants' scores on therapist and significant-other reported TRF instability scales.

Table 5. *Alpha coefficients of the TRF Instability scales (Carlson et al., 2009)*

Domain	Therapist	Significant other
Emotional	.63	.80
Behavioural	.75	.73
Attentional	.66	.73
Interpersonal	.84	.82

TRF: Teacher's Report Form

Table 6. *Descriptive statistics for the TRF Instability scales (Carlson et al., 2009) according to PD traits status*

	TRF (therapist) Mean (SD)		TRF (significant other) Mean (SD)	
	No PD traits (*n*=320)	PD traits (*n*=127)	No PD traits (*n*=289)	PD traits (*n*=125)
Emotional	6.1 (2.6)	6.9 (2.4)	5.6 (3.5)	6.9 (3.4)
Behavioural	3.5 (2.7)	4.9 (2.6)	4.0 (3.1)	4.9 (2.9)
Attentional	4 (2.4)	4.5 (2.3)	3.9 (3.8)	4.8 (3.5)
Interpersonal	2.3 (2.6)	3.4 (2.5)	1.9 (2.6)	2.5 (2.3)

TRF: Teacher's Report Form

Table 7. *Descriptive statistics for the TRF Instability scales (Carlson et al., 2009) according to PD status (PD was operationalised using the alternative, expert-defined criteria)*

	TRF (therapist) Mean (SD)		TRF (significant other) Mean (SD)	
	No PD (*n*=411)	PD (*n*=19)	No PD (*n*=389)	PD (*n*=11)
Emotional	6.3 (2.7)	7.7 (2.6)	6.2 (3.5)	8.1 (3.4)
Behavioural	3.9 (3.4)	7.9 (2.9)	4.3 (3.5)	7.6 (3.3)
Attentional	4.1 (2.8)	4.7 (2.4)	4.2 (3.2)	6.8 (2.9)
Interpersonal	2.6 (3.3)	5.8 (2.8)	3.6 (2.9)	4.2 (2.7)

TRF: Teacher's Report Form

As Table 6 shows, young people with PD traits scored higher on all instability scales. These differences were significant for both therapist [V = .05, F (4, 442) = 5.85, p <.001] and significant-other ratings [V = .03, F (4, 409) = 3.19, p =.013]. For therapist reports, the discriminant analyses that followed revealed one discriminant function, with canonical R^2 .05, and Λ=.95, $\chi^2(4)$= 22.8, p < .001, with 73.6% of original grouped cases

being correctly classified. The scales that contributed most to group separation were the Behavioural (r = .98), Relational (r = .71), and Emotional instability scales (r = .62). For significant-other reports, canonical R^2 was .03, and Λ = .97, $\chi^2(4)$= 12.6, p = .013, with 69.6% of original grouped cases being correctly classified. All scales contributed substantially to group separation, ordered as follows: Emotional (r = .85), Attentional (r = .80), Behavioural (r = .75), and Relational instability (r = .67).

As Table 7 shows, young people with expert-defined PD scored higher on all instability scales. These differences were significant for both therapist [V = .05, F (4, 442) = 5.85, p <.001] and significant-other ratings [V = .05, F (4, 442) = 5.85, p <.001]. For therapist reports, the discriminant analyses that followed revealed one discriminant function, with canonical R^2 .05, and Λ=.94, $\chi^2(4)$= 26.2, p < .001, with 94.9% of original grouped cases being correctly classified. The scales that contributed most to group separation were the Relational (r = .92), and Behavioural (r = .86) instability scales. For significant-other reports, canonical R^2 was .05, and Λ=.96, $\chi^2(4)$= 14.4, p = .006, with 97.3% of original grouped cases being correctly classified. The scales that contributed most to group separation were Behavioural (r = .95), Attentional (r = .77), and Relational instability (r = .51). Again, this alternative PD operationalisation was associated with higher prediction accuracy, but results were overall in high agreement with the findings reported previously when the PD traits definition was used.

Next, a CFA was conducted to test the dimensionality of the TRF instability scales. Overall, the hypothesis that the model fitted the therapist TRF data was rejected [χ^2 (84)= 366.9, p<.001] and the model fit indices fell outside the recommended range: The CFI (.79) and the TLI (.88) were lower than .95 and the RMSEA (.11) was above the recommended cutoff value of.06 (Hu & Bentler, 1998). Furthermore, the WRMR (1.531) exceeded the recommended value of 1. The hypothesis that the model fitted the significant-other TRF data was also rejected [χ^2 (84)= 366.9, p<.001] and the remaining fit

indices fell again outside the recommended range (CFI =.66, TLI =.74, RMSEA =.13, WRMR =1.821).

Dysregulation profile (DP)-PD traits operationalisation

The next set of analyses investigated the discriminatory ability of the dysregulation profile (characterised by co-occurring *T* scores ≥ 70 on the Attention Problems, Aggressive Behaviour, and Anxious/Depressed scales). Initially, χ^2 tests were conducted to test whether there was a difference in the frequency of meeting DP criteria between the groups with and without PD traits.

Results were insignificant for self-reports [$\chi^2(1)=.66$, p =.42, *ns*], but significant for therapist reports [$\chi^2(1)=17.7$, $p < .001$], and significant-other reports [$\chi^2(1)= 5.23$, p=.02], with those meeting DP criteria being more likely to have PD traits. The above analyses were replicated using a cutoff of 60, instead of 70, to explore whether sub-threshold levels of comorbidity were related to PD traits (as in the study by Meyer et al., 2009). When the lower threshold was used, results were significant for all respondent types as follows: for self-reports, $\chi^2(1)=19.66$, $p < .001$, for therapist reports, $\chi^2(1)=28.46$, $p < .001$, and significant-other reports $\chi^2(1)= 5.44$, p=.02, with those meeting DP criteria being more likely to have PD traits.

Next, we used an alternative definition of DP (i.e. the sum of the three scales in question being $\geq$ 180, as defined by Faraone, Althoff, Hudziak, Monuteaux, & Biederman, 2005) to investigate the discriminatory ability of the DP. Results were significant for self-reports [$\chi^2(1)=16.68$, p <.001], therapist reports [$\chi^2(1)=42.97$, $p < .001$], and significant-other reports [$\chi^2(1)= 8.3,3$ p=.02], with those scoring at, or above the 180 cutoff being more likely to have PD traits. At this cutoff, sensitivity and specificity values were .71 and .41 for self-reports, .92 and .24 for therapist reports and .75 and .33 for significant-other reports respectively. To explore the potential usefulness of a different threshold, receiver

operating characteristic (ROC) analysis was conducted to calculate sensitivity values against false positive values (1-specificity) at multiple cutoff scores for the same sum of Attention Problems, Aggressive Behaviour, and Anxious/Depressed scale scores.

We selected a cutoff value that maximised the sum of sensitivity and specificity and minimised overall error. For self-reports, this gave a value of 190, with a sensitivity of .59 and a specificity of .55. For therapist reports, this gave a value of 200, with a sensitivity score of .55 and a specificity score of .67. For significant reports, this gave a value of 194, with a sensitivity of .51 and a specificity of .61. These indices are still quite low, suggesting that the DP sum scale does not have adequate ability to discriminate young people with PD traits.

Dysregulation profile (DP)- PD (expert-defined criteria)

Initially, χ^2 tests were conducted to test whether there was a difference in the frequency of meeting DP criteria between the PD and non-PD groups. Results were significant for self-reports [$\chi^2(1)=7.9$, p =.005], therapist reports [$\chi^2(1)= 46.99$, p<.001], and significant-other reports [$\chi^2(1)=9.59$, p = .002], with those meeting DP criteria being more likely to have PD using the alternative, expert-defined criteria. When the above analyses were replicated using a cutoff of 60, results were again significant for self-reports [$\chi^2(1)=39.05$, p <.001], therapist reports [$\chi^2(1)= 16.30$, p<.001], and significant-other reports [$\chi^2(1)=12.08$, p = .001].

When the alternative definition of DP (i.e. the sum of the three scales in question being $\geq$ 180) was used, results were significant for self-reports [$\chi^2(1)=13.87$, p <.001], therapist reports [$\chi^2(1)=5.46$, p = .019], and significant-other reports [$\chi^2(1)= 8.95$ p=.003], with those scoring at, or above the 180 cutoff being more likely to have PD, operationalised using expert-rated criteria. At this cutoff, sensitivity values were high (.94 for self and therapist reports and .91 for significant-other reports), but specificity values

were unacceptably low (.35 for self-reports, .22 for therapist reports and .32 for significant-other reports).

Using ROC analyses, we selected a cutoff value that maximised the sum of sensitivity and specificity and minimised overall error. For self-reports, this gave a value of 193, with a sensitivity score of .82 and a specificity score of .53. For therapist reports, this gave a value of 208, with a sensitivity score of .58 and a specificity score of .69. For significant-other reports, this gave a value of 197, with a sensitivity of .76 and a specificity of .63. These indices are somewhat improved compared to the findings reported using the PD traits operationalisation, but are still low and imply that the DP sum scale does not have adequate ability to discriminate young people with PD from young people with other mental health problems.

Part II

Item-level YSR data analysis: Kernberg et al.'s (2000) model

Item-level data were available for 830 young people who completed the YSR. This subgroup was overall younger (age *M* = 16.6, *SD* = 2.9), than the remaining participants (*n* = 864, age *M* = 19.3, *SD* = 2.6), which was expected as the YASR was completed by participants older than 18. No other significant differences were detected between this subgroup and the remaining sample.

Table 8 presents the regrouping of the CBCL in terms of PD criteria as suggested by Kernberg et al. (2000). The numbers between parentheses refer to the numbered items in the CBCL. According to this conceptualisation, 57% of the questions on the CBCL reflect particular PDs as they are described in the literature and are presented in DSM-IV, or reveal enduring qualities characteristic of PDs in general.

Table 8. *Kernberg et al.' s (2000, p. 37-39) regrouping of CBCL in terms of PD criteria*

Borderline Personality Disorder	
Argues a lot (3)	Complains of loneliness (12)
Cruelty, bullying, or meanness to others (16)	Deliberately harms self or attempts suicide (18)
Destroys his/ her own things (20)	Feels or complains that no-one loves him/ her (33)
Impulsive or acts without thinking (41)	
Physically attacks other people (57)	
Screams a lot (68)	Sudden changes in mood or feelings (87)
Talks about killing self (91)	Temper tantrums or hot temper (95)
Narcissistic Personality Disorder	
Bragging and boasting (7)	Disobedient at home (22)
Disobedient at school (23)	Feels he/she has to be perfect (32)
Showing off or clowning (74)	
Antisocial Personality Disorder	
Cruel to animals (**15**)*****	Cruelty, bullying, or meanness to others (16)
Destroys things belonging to his/ her family or others (21)	Doesn't seem to feel guilty after misbehaving (26)
Gets in many fights (37)	Impulsive or acts without thinking (41)
Lying or cheating (43)	Runs away from home (67)
Sets fires (72)	Steals at home (81)
Steals outside the home (82)	Threatens people (97)
Truancy, skips school (101)	Vandalism (**106**)*****
Histrionic Personality Disorder	
Demands a lot of attention (19)	Interaction with others is often characterised by inappropriate sexually seductive or provocative behaviour (**73**)*****
Paranoid Personality Disorder	
Easily jealous (27)	Feels others are out to get him/her (34)
Secretive, keeps things to himself/herself (69)	Suspicious (89)
Schizoid Personality Disorder	
Would rather be alone than with others (42)	Doesn't get along with other kids (25)
Strange behaviours (84)	Strange ideas (85)
Withdrawn, doesn't get involved with others (111)	
Avoidant Personality Disorder	
Fears he/she might think or do something bad (31)	Self-conscious or easily embarrassed (71)
	Shy or timid (75)
Dependent Personality Disorder 6	
Acts too young for his/ her age (1)	Clings to adults or is too dependent (11)
Stores up things that he/she does not need (83)	Stubborn, sullen or irritable (86)
Too concerned with meekness and cleanliness (**99**)*****	Whining (**109**)*****
General Personality Disorder Traits	
Feels worthless or inferior (35)	Gets hurt a lot, accident prone (36)
Gets teased a lot (38)	Hangs around with others who get in trouble (39)
Nervous, high-strung, or tense (45)	Not liked by others kids (48)
Too fearful or anxious (50)	Poorly coordinated or clumsy (62)
Prefers being with older kids (63)	Prefers being with younger kids (64)
Sulks a lot (**88**)*****	Talks too much (93)
Teases a lot (94)	Thinks about sex too much (96)
Worries (112)	

*The items with their numbers **in heavy type** were omitted from the CFA testing this model as they are included in the CBCL but in the YSR.

In order to investigate the reliability of the factors/PDs suggested by Kernberg et al. (2000), a CFA was carried out using the YSR item-level data available. As Table 8 shows, six CBCL items that are not shared with the YSR were omitted from the model tested. Furthermore, the factor “histrionic PD” was also omitted as it consisted of only one item found in both the YSR and the CBCL.

The results of the models examined are presented in Table 9. First, an independence model was examined, which assumes that all of the items are uncorrelated and is the standard control in CFA. Second, a nine-factor model was assessed, which presupposes that various items pertain to 8 different PDs and a general PD factor, as in Kernberg et al.’s model of the CBCL (Table 8). As Table 9 shows, this model did not fit the YSR data, and the only factors with acceptable internal consistency were BPD and ASPD (with alphas of .67 and .74 respectively). As a result, we sought to develop a new measurement model.

Item-level YSR data analysis: the YSR ASPD-BPD model

Because ASPD and BPD were the factors that consisted of an adequate number of items and because they were the PD diagnoses the BC clinicians mostly kept in mind when rating PD traits, the 22 items tapping BPD and ASPD became the focus of the remaining analysis, alongside the 14 items considered to reveal general qualities of PDs. Furthermore, item 34 “I feel that others are out to get me” was also included as interpersonal distrust was considered conceptually related to both BPD and ASPD.

At this point, to avoid problems of capitalising on chance fluctuations in the sample, the group was randomly split into two parts, subgroup one and two, each consisting of 415 participants. There were no age or gender differences between the samples (for sample 1, 67% female, age M = 15.7, SD = 4.7 years and for sample 2, 69% female, age M = 15.9, SD = 3.9 years). Using the first sample, a series of Principal Components Analyses (PCA)

were conducted. Varimax rotation was used and a forced two-component solution was specified prior to running the analysis (expected to tap BPD and ASPD).

The Kaiser-Meyer Olkin Measure of Sampling Adequacy was .85 and Bartlett's test of sphericity was highly significant $\chi^2(253)=2504.6$, both confirming the suitability of the data for PCA. Furthermore, all values on the diagonal of the anti-image correlation matrix were well above the bare minimum of .5 (Field, 2005). However, 12 items (26, 36, 38, 63, 64, 67, 72, 81, 82, 93, 96, and 101) with communalities below .2 were discarded, and the remaining 25 items were subjected to another PCA. Two more items (48 and 62) were discarded due to low communalities (<.2). Finally, a third PCA was run with the remaining 23 items and was again subjected to varimax rotation. Table 10 delineates the pattern matrix that emerged from this analysis. The structure matrix revealed a similar configuration.

The final column in Table 10 relates to the CFA and is addressed later. To enhance readability, coefficients that do not exceed .3 are omitted. The component matrix revealed a dominant first factor "BPD" (accounting for 20% of the variance) with 10 out of 23 items having rotated loadings reaching .48 or higher and a second factor "ASPD" (accounting for 13.7% of the variance) with the remaining 13 items having loadings of .45 or higher. Items such as "I deliberately try to hurt or kill myself" (item 18) and "I feel lonely" (item 12) loaded meaningfully on the BPD component whereas items such as "I argue a lot" (item 3) and "I threaten to hurt people" (item 97) loaded meaningfully on the ASPD component.

Table 9. *Goodness-of-Fit Indices generated by the Confirmatory Factor Analysis (CFA) of the YSR items*

Model	χ^2	df	CFI	TLI	RMSEA	WRMR
Independence model	4813.4	131				
Kernberg et al.'s model	2947.8	317	.55	.69	.10	2.54
Two-factor oblique model						
Subgroup 1 (*n* = 415)	345.7	147	.85	.87	.094	1.57
Subgroup 2 (*n* = 415)	432.1	147	.83	.86	.097	1.72
Males (*n* = 234)*	225.4	65	.84	.87	.103	1.44
Females (n=568)*	533.8	91	.83	.89	.093	1.83
Whole sample (*n*=830)	843.7	92	.82	.87	.099	2.31
Two-factor orthogonal model (*n*=830)	714.9	62	.79	.80	.113	2.74
One-factor model i.e. ASPD-BPD items combined (*n*=830)	2769	83	.35	.49	.197	4.52

Notes. CFI = Comparative Fit Index, TLI = Tucker-Lewis Index, RMSEA = Root Mean Square Error of Approximation Index, WRMR= Weighted Root Mean Square Residual.
The $\chi^2$and degrees of freedom are adjusted to obtain a correct *p*-value with weighted least-squares means and variance adjusted estimation (WLSMV).
*Frequencies do not add up to total due to missing data

Table 10. *Factor Pattern Matrix Rotated to the Varimax Criterion*

YSR items		BPD	ASPD	h^2	CFA
Item 12.	I feel lonely	.739		.552	.730
Item 18.	I deliberately try to hurt or kill myself	.520		.350	.710
Item 33.	I feel that no one loves me	.626		.401	.673
Item 34.	I feel that others are out to get me	.519		.351	.616
Item 35.	I feel worthless or inferior	.781		.610	.791
Item 45.	I am nervous or tense	.698		.487	.732
Item 50.	I am too fearful or anxious	.679		.472	.690
Item 87.	My mood or feelings change suddenly	.478		.314	.571
Item 91.	I think about killing myself	.635		.440	.749
Item 112.	I worry a lot	.675		.470	.709
Item **3***.	I argue a lot		.568	.329	.629
Item **16***.	I am mean to others		.512	.279	.597
Item 20.	I destroy my own things		.486	.296	.625
Item **21***.	I destroy things belonging to others		.471	.228	.612
Item **37***.	I get in many fights		.621	.407	.609
Item **39***.	I hang around with kids who get in trouble		.494	.263	.439
Item 41.	I act without stopping to think		.504	.261	.519
Item **43***.	I lie or cheat		.451	.204	.459
Item **57***.	I physically attack people		.640	.409	.739
Item 68.	I scream a lot		.455	.257	.570
Item 94.	I tease others a lot		.471	.222	.490
Item **95***.	I have a hot temper		.612	.400	.721
Item **97***.	I threaten to hurt people		.671	.450	.752
Trace		0.87	0.60	1.465	
% of variance		20%	13.7%	33.7%	

Note. *N* = 415. Percentage variance is post-rotation. Because here there were 23 measured variables, percentage of variance is trace divided by 23 times 100 (or trace times 23). The last column presents the standardised coefficients that emerged from the confirmatory factor analysis (CFA). h^2 =communality coefficient; YSR = Youth Self-Report

*The items with their numbers **in heavy type** are included in the Antisocial Personality Problems scale of the Young Adult Self-Report.

To investigate the reliability of these factors, a CFA was carried out using MPlus with the second sample. The standardised coefficients associated with each item are presented in the last column of Table 10. All of these coefficients exceed .4, providing initial support for the efficacy of the model. The efficacy of the derived model was further examined for the whole group, and also across gender. As Table 9 shows, the model's fit indices were improved, compared to Kernberg et al.'s (2000) model, but also diverted from the recommended criteria of a good measurement model. In addition, the χ^2 tests were all significant, indicating that the null hypothesis - that the model does fit the data - should be rejected. However, it is worth pointing out that a non-significant χ^2 is rarely obtained when sample sizes are large (Bentler & Bonett, 1980; Joreskog, 1981).

The relevant orthogonal model, which does not allow the BPD and ASPD factors to correlate, and the one-factor model, which presupposes that all of the items pertain to the same factor, were also scrutinised but provided a poorer fit for the data compared to the oblique model, which assumes that BPD and ASPD are different, but related constructs (Table 9).

On the whole, although the fit indices surpassed the criteria of an acceptable two factor oblique model, no definite criteria exist to determine precise cutoffs, and "interpretation of fit indices has to take into account a number of measures as well as the nature of the data and the model under examination" (Heubeck, 2000, p. 443). Given the theoretical coherence of the emerged factors, their good internal consistency (see below), and the fact that this model fitted our data better than both Kernberg et al.'s (2000) and Carlson et al.'s (2009) proposed item groupings, we investigated the model further.

Distribution of scores

Next, we aggregated the responses given to the 10 BPD items into total scores for the whole sample. The mean was 9.3 (*SD* = 4.7), the median 9, the mode 6, and the range

0 - 20. We next aggregated the responses given to the 13 ASPD items. The mean was 7.1 (*SD* = 4.4), the median 6, the mode 5, and the range 0 - 23. Both distributions were positively skewed. The correlation between the BPD and ASPD factors was *r* (831)=.22, *p*< .001.

Reliability

The next set of analyses investigates the reliability of the two scales. The alpha for the BPD scale was high at .85 and for the ASPD scale was also high at .80. Corrected item-total correlation coefficients ranged from .415 (item 87) to .685 (item 35) for the BPD scale, and from .360 (items 39 and 43) to .542 (item 97) for the ASPD scale.

Validity

As Table 11 shows, the PD (traits) group tended to have higher scores on both the BPD and the ASPD scales, although the significance of these differences depended on the PD operationalisation used. Furthermore, females scored higher on the BPD scale, whereas males scored higher on the ASPD scale. Furthermore, both scales correlated negatively with the GAF scale, for BPD, *r(*786)=-.26, *p*<.001, and for ASPD, *r*(786)=-.13, *p*<.001. In addition, age correlated positively with BPD, and negatively with ASPD *r*(805)=-.32, *p*<.001, and for ASPD, *r*(805)=-.21, *p*<.001. These results are overall in support of the construct validity of these scales.

Finally, ROC curves were employed to investigate the capacity of the BPD and ASPD factors to differentiate between the PD and non-PD (traits) groups. For the BPD factor, the value that maximised the sum of sensitivity and specificity in discriminating between the PD-non PD traits groups was 9.5, and yielded a sensitivity score of .60 and a specificity score of .55. For the ASPD factor, the same cutoff (5.5) yielded a sensitivity score of .59 and a specificity score of .47. Using the expert-defined PD criteria, a BPD

score of 10.5 yielded a sensitivity score of .55 and a specificity score of .58. For the ASPD factor, the 8.5 cutoff yielded a sensitivity score of .73 and a specificity score of .71.

Table 11. *Differences between young people with and without PD (traits) and gender differences on the YSR BPD and ASPD scales*

YSR PD scales	No PD traits (*n*=447)	PD traits (*n*=182)	*t* test
BPD	9.1 (4.5)	11(4.6)	*t*(627)=4.8***
ASPD	6.7(4.3)	7.5(4.4)	*t*(627)=1.9, *p*=.056, ns.
	No PD (*n*=578)	PD (*n*=22)	(using expert defined criteria)
BPD	9.6 (4.6)	11.1 (3.9)	*t*(598)=1.5,*p*=.127, ns.
ASPD	6.9 (4.3)	10.3 (4.5)	*t*(598)=3.6***
	Males (*n*= 236)	Females (*n*=570)	
BPD	8(4.8)	10 (4.4)	*t*(804)=5.4***
ASPD	7.5(4.1)	6.8 (4.5)	*t*(804)=2.2*

*p<.05 **p<.01 ***p<.001

Discussion

Overview of findings

At the beginning of this research we set out to explore whether four ASEBA forms, namely the YSR, the TRF, the YASR and the YABCL could be used to identify PD in young people. Given that these forms are widely used for routine outcome monitoring, and because they are a research standard in the dimensional approach to psychiatric assessment, we wanted to know whether young people with PD traits scored differently on these forms compared to young people with other mental health problems. Furthermore we wanted to identify the scales (or combination of scales or items) that discriminated mostly between the PD-non PD groups. To address these questions, a variety of statistical analyses were conducted. How are we to interpret the results?

Our findings showed that overall, young people considered to have at least some PD traits/typical PD symptoms scored higher on all self-report ASEBA scales, compared to

their peers without PD traits and symptoms. However, the effect size of these differences was small overall, and rarely greater than 3 *T* scores. The scales that contributed most to group separation were Aggressive Behaviour, Delinquent Behaviour, and Thought Problems. The latter two syndromes remained significant predictors of PD after controlling for other known predictors (such as gender and mood disorder). Furthermore, the Anxious/Depressed syndrome was also a significant predictor, but a negative one (with lower scores on this scale predicting PD).

These results were largely replicated when therapist reports were used, with somewhat greater effect sizes, reflecting the impact of shared method variance. In addition, the Somatic Symptoms syndrome was also a significant predictor, but a negative one (with higher scores reducing the likelihood of PD). The findings concerning significant-others' reports were overall consistent with the above. The Thought and Attention Problems, Delinquent and Aggressive Behaviour and Withdrawn syndromes contributed most to the PD-non PD differentiation, although again the effect sizes were quite small.

These results are meaningful and largely in line with the criteria of ASPD and BPD. In particular, Delinquent Behaviour is a defining feature of ASPD, and the Aggressive Behaviour syndrome is also a defining feature of ASPD. This syndrome consists of items pertaining to both physical aggression (e.g. item 37 about getting in many fights and item 57 about physically attacking people) and relational aggression (e.g. item 16 about being mean to others and item 94 about teasing others a lot). The latter type has been found to also correlate with borderline personality traits in middle childhood (Crick et al., 2005) and late adolescence (Ostrov & Houston, 2008; Werner & Crick, 1999). However, the items that comprised the ASPD scale pertain to both aggressive and non-aggressive conduct problems and it was not possible to test whether this scale could discriminate between these two types of conduct problems.

The Thought Problems scale includes an item about self-harm and items concerning unusual thoughts, behaviours and sensory experiences. These are in keeping with DSM-IV-TR (APA, 2000) criteria of BPD (in particular those regarding self-harm and transient psychotic experiences). In addition, the Attention Problems syndrome includes items regarding impulsivity and daydreaming (which conceptually relates to dissociation), that are also consistent with DSM-IV criteria of BPD. The Withdrawn scale consists of items related to social withdrawal and under-activity that are more relevant to Cluster A PDs. Moreover, the mild, inverse association between PD and the Anxious/Depressed and Somatic Complaints syndromes is more in line with the phenomenology of ASPD.

Because no information was available regarding the specific PD type of which participants had problems and traits, the above observations cannot be tested empirically and remain tentative. However, these findings are explicable within the wider literature that suggests that externalising problems correlate positively with juvenile ASPD (e.g. Barnow, Lucht, & Freyberger, 2005; Dolan & Rennie, 2006, 2007; Fite, Greening, & Stoppelbein, 2008; Sharp, Mosko, Chang, & Ha, 2011; Westen, Shedler, Durrett, Glass, & Martens, 2003), and BPD (e.g. Arens, Grabe, Spitzer, & Barnow, 2011; Burnette, South, & Reppucci, 2007; Rogosch & Cicchetti, 2005; Sharp, Mosko, et al., 2011; Sharp, Pane, et al., 2011; Sharp et al., 2012; Zelkowitz et al., 2007). Thought Problems have also been associated with juvenile schizotypal PD (Wickline, Nowicki, Bollini, & Walker, 2012), ASPD (Lexcen, Vincent, & Grisso, 2004) and BPD (Underwood, Beron, & Rosen, 2011; Zelkowitz et al., 2007).

Although a large percentage of cases were correctly classified using the function derived from the discriminant analysis, this was partly the result of basing prior probabilities on the observed group sizes; the modest canonical R^2 obtained (below 10% in most cases) suggests that the ASEBA syndromes contributed only minimally to the differentiation between the PD and the non-PD group. Furthermore, although some ASEBA scales remained significant predictors of PD after controlling for the influence of

other factors, effect sizes were small, and it would have been unlikely to detect them with a smaller sample.

The next set of analyses concerned the use of the TRF-instability scales. The internal consistency indices of these scales were found to be comparable to those reported by Carlson et al. (2009). In addition, the group with PD (traits) tended to have higher scores on all these scales than their non-PD peers. The behavioural instability scale in particular was the one that most consistently differentiated the PD and non-PD groups regardless of informant type. However, the four-domain grouping of the items was not supported in our data, as the CFA fit indices fell outside the recommended range.

Subsequently, we investigated the use of the DP. Young people meeting DP criteria (regardless of informant type) were more likely to have PD (traits). However, when the discriminatory ability of the DP profile was investigated with ROC analysis (using the sum of Attention Problems, Aggressive Behaviour, and Anxious/Depressed scale scores) no cutoff point with adequate sensitivity and specificity could be identified. This is probably because in our data, only the Aggressive Behaviour syndrome was consistently found to differentiate between the PD and non-PD groups.

The second part of the results presented an item-level analysis of the YSR, in an effort to a) test Kernberg et al.'s (2000) hypothesis about the PD grouping of CBCL items and b) develop alternative, empirically supported PD-oriented scales with adequate reliability and validity. Kernberg et al.'s measurement model was not supported by the CFA, possibly because most PD factors (apart from ASPD and BPD) consisted of a limited number of items. Therefore, we subsequently developed an ASPD and a BPD scale that was based on (but not identical to) Kernberg et al.'s ASPD and BPD factors.

Both new scales demonstrated good internal consistency. In terms of content validity, the majority of the items are consistent with DSM criteria of ASPD and BPD. Furthermore, many of the items are in accordance with the core PD criteria identified by Geiger and Crick (2001): Namely, they reflect a presence of negative self-view (e.g. item

35 “I feel worthless or inferior”), intense, unstable, and inappropriate emotion (e.g. item 87 “My mood or feelings change suddenly”), a hostile world view (e.g. item 34 “I feel that others are out to get me”) and rigidity or impulsivity (e.g. item 4 “I act without stopping to think”). However, not all DSM-IV criteria were represented by the items selected (e.g. chronic feelings of emptiness in relation to BPD, and lack of remorse in relation to ASPD). Some items are also broadly consistent with the DSM-V (APA, 2012) proposed PD criteria concerning impairment in self-identity and interpersonal functioning. Furthermore, most ASPD items relate to antisocial and aggressive *behaviours*, whereas most BPD items are about *psychological states,* in accordance with the essential features of each PD diagnosis.

In terms of concurrent (criterion) validity, the PD (traits) group scored higher on both scales, although this difference did not always reach statistical significance. In terms of construct validity, the negative correlations of these scales with clinicians’ ratings of general functioning are meaningful and unaffected by shared method variance. This association is in accordance with evidence from numerous studies suggesting that adolescent PD, especially BPD and ASPD, is associated with poor clinical and psychosocial functioning and characterised by high comorbidity rates with Axis I disorders (Sharp, Pane, et al., 2011).

In terms of convergent (construct) validity, the gender differences found are in agreement with a large body of evidence documenting that BPD is more prevalent among females, whereas ASPD is more prevalent among males (Morey, Alexander & Boggs, 2005), and this is also indicated by DSM-IV TR (APA, 2000). The correlation between the ASPD and BPD scales derived from the YSR is also broadly consistent with the conceptual overlap and comorbidity estimates regarding these disorders in clinical samples (Paris, 1997). The positive association between age and BPD and the negative association between age and ASPD is harder to interpret but it may relate to the age range of our sample (12-25).

Despite these positive attributes, the discriminatory ability of both the BPD and the ASPD scales was low and inadequate for clinical use. Another weakness of these scales relates to the finding that the measurement model was not supported by the CFA. If cross-loadings had been allowed, the model would have resulted in marginally good fit indices. However, when conducting a CFA, the analysis should not be governed by the fit indices of the model alone. There are also conceptual factors to consider, and parameters such as the factor loading for each observed variable are also important (Farrell & Rudd, 2009). In any case, no other model achieved a good fit to the study data. In fact, the BPD-ASPD model had the fit indices closer to the recommended levels compared to both Carlson et al.'s (2009) and Kernberg et al.'s (2000) models.

Potential Objections and Limitations

One might raise two primary objections to this study. First, the reliability of our dependent variable (clinician-based PD diagnosis) was not established, and in addition this diagnostic category was heterogeneous as it included apart from PD, other clinical problems. This confound is obviously a significant methodological problem. More generally, there is evidence that in the absence of psychometrically sound measures, clinician-based diagnoses may be non-systematic and prone to bias (Dutra, Campbell, & Westen, 2004).

Although an alternative, more robust method of operationalising PD was used, this definition was also limited in that it only focused on a limited number of PD characteristics (those on which a high level of expert agreement was reached), leaving many important other features (e.g. interpersonal problems) unaddressed. Furthermore, because our approach was conservative in that participants had to have all four problems to be considered diagnosable with PD, the sensitivity of this operationalisation was low.

The second primary objection relates to the absence of any specific PD type data. As a result, no differentiation could be made between participants with various PD traits

and characteristics. This means that, in theory, the PD group could have consisted of young people with very different presentations. For instance, people with obsessive compulsive PD can be on the polar opposite of people with ASPD or BPD with respect to impulsivity. This high degree of heterogeneity may have confounded our findings in relation to the discriminatory ability of the various ASEBA scales, including the ASPD and BPD scales developed in this study. This is because, for instance, whilst items from the Aggressive Behaviour syndrome may differentiate well between those with and without ASPD, the same scale will be pretty ineffective in differentiating between people with OCPD or other Cluster C disorders. However, although the PD type was not specified, most clinicians used the ASPD and BPD criteria to rate the severity of any PD-related problems.

A few other limitations of this study should be noted. The correlational and cross-sectional nature of the data did not allow for the predictive validity of the ASEBA to be assessed. An additional limitation is the lack of data concerning the test-retest reliability of the ASPD and BPD scales. Furthermore, the results concerning the use of the YABCL and the YASR are somewhat dated as these forms have been now replaced by the ABCL and the ASR. However, the overlap between the young adult and adult forms is extensive and our findings should be replicated with the adult forms (and a young adult population). In addition, the use of an alternative, expert-defined operationalisation of PD resulted in tautology issues and probably inflated associations with ASEBA scales due to overlapping item content (this concerns mainly ASEBA items that enquire about self-harm, substance misuse and antisocial behaviour).

Despite these limitations, a few strengths of this study should be acknowledged. Firstly, we used a unique community and clinical practice-based dataset with a large percentage of those presenting being included in the analysis (as opposed to convenience samples gathered from schools or colleges, or selective samples from university clinics).

Secondly, the large sample size facilitated the application of rigorous statistical techniques and allowed the testing of complex models. This for instance enabled us to investigate the ability of various ASEBA scales to statistically predict PD after controlling for other factors such as gender and mood disorder. Thirdly, there can be no argument that the multi-method measurement strategy is superior to the use of a single observer, and the data from multiple observers enabled the study of the validity of various ASEBA scales above and beyond shared method variance. Fourthly, the availability of a dataset that included all ASEBA syndrome scales enabled a thorough investigation and comparison of these scales within the same sample.

Clinical implications

The ASEBA showed potential for providing useful clinical information about PD-related problems in young people. Consequently, it may be useful for detecting clinical and sub-clinical levels of PD in community mental health settings. However, the findings of this study should be considered preliminary in the absence of a reliable PD criterion. This may partially explain why no ASEBA syndrome (or set of syndromes/items) had adequate discriminatory ability to be recommended as a reliable and valid screening measure for PD.

Diagnostic accuracy in mental health is crucial as it enables practitioners to plan suitable interventions and prevent possible iatrogenic harm. Nevertheless, this study was unable to detect an ASEBA scale with a cutoff point with adequate sensitivity and specificity to recommend as a screening tool of emergent PD. However, even without an established diagnostic cutoff point, the examination of the BPD and ASPD scales could be useful for providing clinicians with information about the extent to which the young people they are assessing have symptoms indicative of emerging PD, with a view to case formulation and active treatment planning. Furthermore, these results also have significant

implications for the use of the ASEBA in longitudinal studies to help illuminate the developmental precursors of PD.

Directions for future research

As all ASEBA forms have been standardised within very large populations, norms for the ASPD and BPD scales could be obtained to develop clinical cutoff scores to aid clinical decision making. Most importantly, further research should use well-established PD measures to evaluate the validity of various syndrome scales and the newly developed BPD and ASPD scales. There is a need to look at the clinical usefulness of these scales for identifying adolescents diagnosed with PD in accordance with standardised clinical interviews, and also to explore the concurrent validity of these scales with well-validated questionnaire measures of PD.

Moreover, future research is needed to investigate the reliability of these scales over time. Additional longitudinal research is required to examine the extent to which the ASEBA can be used to assess clinical outcome (including attrition) as well as treatment effect sizes in this and other populations. Future prospective research is also needed to examine which child and adolescent ASEBA scales predict PD in adulthood.

Future studies will need to include large samples so that the effects of relevant demographic (e.g. age and gender) and clinical (e.g. depression) variables can be controlled for. The examination of the derived YSR BPD and ASPD scales with other informant types (e.g. parents) and ASEBA forms (e.g. the CBCL) will be informative as to whether these scales are applicable to younger children and non-self-report measures, and will also provide data regarding cross-informant correlations. In addition, when assessing PD, it is important to obtain information from sources other than the client (Kernberg et al., 2000).

Finally, advanced statistical analyses are expected to further the examination of these scales as they have many advantages over traditional psychometric practices (Reise

& Waller, 2009). For instance, latent class analysis can identify clusters of participants with statistical elevations on relevant scales without imposing arbitrary cutoff points (Althoff et al., 2012). Likewise, item response theory (Embretson & Reise, 2000, cited in Reise & Waller, 2009) emphasises the study of response processes and the meaning of latent traits, instead of establishing a network of external correlations with other scales (Reise & Waller, 2009).

Conclusion

Linking nosological and statistical paradigms in the study of emerging PD holds promise but measures need to be empirically based and clinically relevant. The establishment of robust psychometric properties is a much-needed requirement before the identification of PD features can be integrated into regular screening of psychiatric problems in child and adolescent mental health services (Sharp, Mosko, et al., 2011). The field of PD assessment and diagnosis is still evolving and error prone, even in adults (Clark, 2007), but more accurate diagnostic procedures will likely be developed as the field continues to mature. Although the results of this study indicated that none of the ASEBA forms or scales can be used as efficient PD diagnostic tools, it is possible that with further research we will ultimately be able to benefit from the scientific history and popularity of the ASEBA to develop an effective PD screening tool for young people. A high degree of psychometric sophistication will be unquestionably required, and despite its several limitations, this study has provided initial insights into this direction.

References

Achenbach, T. M. (1991). *Integrative Guide for the 1991 CBCL/4-18, YSR, and TRF profiles*. Burlington, VT: University of Vermont, Department of Psychiatry.

Achenbach, T.M. (1997*).Manual for the Young Adult Self-Report and Young Adult Behavior Checklist*. Burlington, Vermont: University of Vermont, Department of Psychiatry.

Achenbach, T. M. (2009). *The Achenbach System of Empirically Based Assessment (ASEBA): Development, Findings, Theory, and Applications*. Burlington, VT: University of Vermont Research Center for Children, Youth and Families.

Achenbach, T. M., Bernstein, A., & Dumenci, L. (2005). DSM-Oriented Scales and Statistically Based Syndromes for Ages 18 to 59: Linking Taxonomic Paradigms to Facilitate Multitaxonomic Approaches, *Journal of Personality Assessment, 84*(1), 49-63.

Achenbach, T. M., Dumenci, L., & Rescorla, L. A. (2001). *Ratings of Relations Between DSM-IV Diagnostic Categories and Items of the CBCL/6-18, TRF, and YSR*. Burlington: University of Vermont, Research Center for Children, Youth, and Families. Available at www.ASEBA.org.

Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin, 101*, 213-232.

Achenbach, T. M., & Rescorla, L. A. (2003). *Manual for the ASEBA Adult Forms & Profiles.* Burlington: University of Vermont, Research Center for Children, Youth, and Families.

Althoff, R.R., Rettew, D.C., Faraone, S.V., Boomsma, D.I., & Hudziak, J.J. (2006). Latent class analysis shows strong heritability of the child behavior checklist-juvenile bipolar phenotype. *Biological Psychiatry, 60*(9), 903-11.

American Psychiatric Association (1980). *Diagnostic and Statistical Manual of Mental Disorders* (*3rd edition*). Washington, DC: American Psychiatric Association.

American Psychiatric Association (1987). *Diagnostic and Statistical Manual of Mental Disorders* (*3rd edition, text revision*). Washington, DC: American Psychiatric Association.

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

American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders* (*4th edition, text revision*). Washington, DC: American Psychiatric Association.

American Psychiatric Association (2012). Personality and Personality Disorders. http://www.dsm5.org/PROPOSEDREVISIONS/Pages/PersonalityandPersonalityDisorders.aspx

Arens, E. A., Grabe, H. J., Spitzer, C., & Barnow, S. (2011). Testing the biosocial model of borderline personality disorder: Results of a prospective 5-year longitudinal study. *Personality and Mental Health, 5*(1), 29-42.

Bargmann, R.E. (1970). Interpretation and use of a generalised discriminant function. In R.C. Bose I. M. Chakravarti, P. C. Mahalanobis, C. R. Rao & K. J. C. Smith (Eds.) *Essays in probability and statistics*, Univ. North Carolina Press, Chapel Hill, pp. 35-60.

Barnow, S., Lucht, M., & Freyberger, H. (2005). Correlates of aggressive and delinquent conduct problems in adolescence. *Aggressive Behavior, 31*(1), 24-39.

Baruch, G. (1995). Evaluating the outcome of a community based psychoanalytic psychotherapy service for young people between 12 and 25 years old: Work in progress. *Psychoanalytic Psychotherapy, 9,* 243–267.

Baruch, G., & Fearon, P. (2001). The routine evaluation of mental health outcome at a community-based psychotherapy centre for young people. In G. Baruch (Ed.), *Community-Based Psychotherapy with Young People*. Hove: Brunner Routledge.

Baruch, G., Fearon, P., & Gerber, A. (1998). Evaluating the outcome of a community-based psychoanalytic psychotherapy service for young people: One-year repeated follow-up. In R. Davenhill & M. Patrick (Eds.). *Rethinking Clinical Audit: The Case of Psychotherapy Services in the NHS*. London: Routledge.

Baruch, G., Gerber, A., & Fearon, P. (1998). Adolescents who drop out of psychotherapy at a community-based psychotherapy centre: A preliminary investigation of the characteristics of early dropouts, late dropouts and those who continue treatment. *British Journal of Medical Psychology, 71,* 233-245.

Baruch, G., & Vrouva, I. (2010). Collecting routine outcome data in a psychotherapy

community clinic for young people: Findings from an ongoing study. *Child and Adolescent Mental Health, 15* (1), 30-36.

Baruch, G., Vrouva, I., & Wells, C. (2011). Outcome findings from a parenting programme for young people with conduct problems. *Child and Adolescent Mental Health, 16,* 47-54.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*(2), 238-246.

Bentler, P.M., & Bonett, D.G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88*, 588-606.

Bérubé, R. L., & Achenbach, T. M. (2007). *Bibliography of Published Studies Using the ASEBA*. Burlington, VT: University of Vermont, Research Center for Children, Youth & Families.

Breckler, S. J. (1990). Applications of covariance structure modeling in psychology: Cause for concern? *Psychological Bulletin, 107,* 260–273.

Burnette, M. L., South, S. C., & Reppucci, N. D. (2007). Cluster B personality pathology in incarcerated girls: Structure, co-morbidity, and aggression. *Journal of Personality Disorders, 21*(3), 262-272.

Carlson, E. A., Egeland, B., & Sroufe, L.A. (2009). A prospective investigation of the development of borderline personality symptoms. *Development and Psychopathology, 21* (4), 1311-1334.

Clark, L. A. (2007). Assessment and diagnosis of personality disorder: Perennial issues and an emerging reconceptualization. *Annual Review of Psychology, 58,* 227-57.

Cohen, J. (1977). *Statistical Power Analysis for the Behavioural Sciences* (rev. ed.). New York: Academic Press.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155-159.

Cohen, P. (2008). Child development and personality disorder. *Psychiatric Clinics of North America, 31*(3), 477-493.

Cook, R.D., & Weisberg, S. (1982). *Residuals and Influence in Regression*. New York: Chapman & Hall.

Crawford, T.N., Cohen, P., First, M.B., Skodol, A.E., Johnson, J.G., & Kasen, S. (2008). Comorbid Axis I and Axis II disorders in early adolescence. *Archives of General Psychiatry, 65*(6), 641-648.

Crick, N. R., Murray-Close, D., & Woods, K. (2005). Borderline personality features in childhood: A short-term longitudinal study. *Development and Psychopathology, 17*(4), 1051-1070.

Cudeck, R., & Browne, M. W. (1983). Cross-validation of covariance structures. *Multivariate Behavioral Research, 18,* 147–167.

Cronbach, L. J., & Shavelson, R. J. (2004). My current thoughts on coefficient alpha and successor procedures. *Educational and Psychological Measurement, 64*(3), 391–418.

De Clercq, B., & De Fruyt, F. (2007). Childhood antecedents of personality disorder. *Current Opinion in Psychiatry, 20*(1), 57-61.

De Clercq, B., De Fruyt, F., Van Leeuwen, K., & Mervielde, I. (2006). The structure of maladaptive personality traits in childhood: A step toward an integrative developmental perspective for DSM-V. *Journal of Abnormal Psychology,115*(4), 639-657.

De Groot, A., Koot, H.M., & Verhulst, F.C. (1994). Cross-cultural generalizability of the Child Behavior Checklist cross-informant syndromes. *Psychological Assessments, 6,* 225–230.

Dolan, M. C. & Rennie, C. E. (2006). Reliability and validity of the psychopathy checklist: Youth version in a UK sample of conduct disordered boys. *Personality and Individual Differences, 40* (1), 65-75.

Dolan, M. C., & Rennie, C. E. (2007). The relationship between psychopathic traits measured by the Youth Psychopathic Trait Inventory and psychopathology in a UK sample of conduct disordered boys*. Journal of Adolescence, 30*(4), 601-611.

Dutra, L., Campbell, L., & Westen, D. (2004). Quantifying clinical judgment in the assessment of adolescent psychopathology: Reliability, validity, and factor structure of the Child Behavior Checklist for clinician report. *Journal of Clinical Psychology, 60*(1), 65-85.

Eggum, N. D., Eisenberg, N., Spinrad, T. L., Valiente, C., Edwards, A., Kupfer, A. S., & Reiser. M. (2009). Predictors of withdrawal: Possible precursors of avoidant personality disorder. *Development and Psychopathology, 21*(3), 815-838.

Embretson, S.E., Reise, S.P. (2000). *Item Response Theory for Psychologists*. Mahwah, NJ: Erlbaum.

Faraone, S. V., Althoff, R. R., Hudziak, J. J., Monuteaux, M., & Biederman, J. (2005). The CBCL predicts DSM bipolar disorder in children: a receiver operating characteristic curve analysis. *Bipolar Disorders*, 7(6), 518-524.

Farrell, A.M., & Rudd, J.M. (2009). "Factor Analysis and Discriminant Validity: A Brief Review of Some Practical Issues" Australia-New Zealand Marketing Academy Conference (ANZMAC), December, Melbourne, Australia.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods*, *39*(2), 175–191.

Ferdinand, R.F., & Verhulst, F.C. (1994). The prediction of poor outcome in young adults: comparison of the Young Adult Self-Report, the General Health Questionnaire, and the Symptom Checklist. *Acta Psychiatrica Scandinavica, 89,* 405-410.

Field, A. (2005). *Discovering statistics using SPSS.* (2nd ed.). London: Sage.

Fite, P. J., Greening, L., & Stoppelbein, L. (2008). Relation between parenting stress and psychopathic traits among children. *Behavioral Sciences & the Law, 26,* 239-248.

Fonagy, P., & Higgitt, A. (1989). Evaluating the performance of departments of psychotherapy. *Psychoanalytic Psychotherapy, 4,* 121–153.

Gagne, P. & Hancock, G. R. (2006). Measurement model quality, sample size, and solution propriety in confirmatory factor models. *Multivariate Behavioral Research, 41*(1), 65-83.

Geiger, T., & Crick, N. R. (2001). A developmental psychopathology perspective on vulnerability to personality disorders. In R. Ingram & J. M. Price (Eds.), *Vulnerability to Psychopathology: Risk Across the Life Span* (pp.57-102). New York: Guilford.

Halperin, J. M., Rucklidge, J. J., Powers, R. L., Miller, C. J., & Newcorn, J. H. (2011). Childhood CBCL bipolar profile and adolescent/young adult personality disorders: A 9-year follow-up. *Journal of Affective Disorders, 130*, 155-161.

Heubeck. B. G. (2000). Cross-cultural generalizability of CBCL syndromes across three continents: From the USA and Holland to Australia. *Journal of Abnormal Child Psychology, 28*, 439-450.

Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to under parameterized model misspecification. *Psychological Methods, 3*, 424-453.

Hudziak, J.J., Althoff, R.R., Derks, E.M., Faraone, S.V., Boomsma, D.I. (2005). Prevalence and genetic architecture of Child Behavior Checklist-juvenile bipolar disorder. *Biological Psychiatry, 58*(7), 562-8.

Joreskog, S.G. (1981). Analysis of covariance structures. *Scandinavian Journal of Statistics, 8*, 65-92.

Kernberg, P., Weiner, A., & Bardenstein, K. (2000). *Personality Disorders in Children and Adolescents.* New York: Basic books.

Kolko, D. J., & Kazdin, A. E. (1993). Emotional/behavioral problems in clinic and nonclinic children: Correspondence among child, parent, and teacher reports. *Journal of Child Psychology and Psychiatry, 34,* 991-1006.

Lexcen, F. J., Vincent, G. M., & Grisso, T. (2004). Validity and structure of a self-report measure of youth psychopathy. *Behavioral Sciences & the Law, 22,* 69-84.

MacCallum, R. C., Roznowski, M., & Necowitz, L. B. (1992). Model modifications in covariance structure analysis: The problem of capitalization on chance. *Psychological Bulletin, 111,* 490–504.

Meyer, S. E., Carlson, G. A., Youngstrom, E., Ronsaville, D. S., Martinez, P. E., Gold, P. W., et al. (2009). Long-term outcomes of youth who manifested the CBCL-Pediatric Bipolar Disorder phenotype during childhood and/or adolescence. *Journal of Affective Disorders, 113*(3), 227-235.

Morey, L.C., Alexander, G.M., & Boggs, C. (2005). Gender and personality disorder. In J. M. Oldham, A. E. Skodol & D. S. Bender (Eds.), *The American Psychiatric Publishing Textbook of Personality Disorders* (pp. 541-559). Washington, DC: American Psychiatric Press.

Muthén, B., Du Toit, S. H. C., & Spisic, D. (1997). *Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with categorical and continuous outcomes*. Unpublished technical report.

Muthén, L. K., & Muthén, B. O. (2007). *MPlus User's Guide* (5th ed.). Los Angeles, CA: Author.

National Institute for Mental Health in England (NIMHE) (2003, 23 January). *Personality Disorder: No Longer a Diagnosis of Exclusion*. Retrieved 20 May 2012 from the Department of Health website: http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_4009546

Ostrov, J. M., & Houston, R. J. (2008). The utility of forms and functions of aggression in emerging adulthood: Association with personality disorder symptomatology. *Journal of Youth and Adolescence, 37,* 1147–1158.

Paris, J. (1997). Antisocial and borderline personality disorders: two separate diagnoses or two aspects of the same psychopathology? *Comprehensive Psychiatry, 38,* 237–242.

Phares, V., Comas, B. E., & Howell, D. C. (1989). Perspectives on child behavior problems: Comparisons of children's self-reports with parent and teacher reports. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 1,* 68-71.

Reise, S.P., & Waller, N.G. (2009). Item Response Theory and Clinical Measurement. *Annual Review of Clinical Psychology, 5*, 27–48.

Rogosch, F. A., & Cicchetti, D. (2005). Child maltreatment, attention networks, and potential precursors to borderline personality disorder. *Development and Psychopathology, 17* (4), 1071-1089.

Seagrave, D., & Grisso, T. (2002). Adolescent development and the measurement of juvenile of psychopathy. *Law and Human Behavior*, *26*, 219-239.

Sharp, C., Ha, C., Michonski, J., Venta, A., & Carbone, C. (2012). Borderline personality disorder in adolescents: Evidence in support of the Childhood Interview for DSM-IV Borderline Personality Disorder in a sample of adolescent inpatients. *Comprehensive Psychiatry, 53*(6), 765-774.

Sharp, C., Mosko, O., Chang, B., & Ha, C. (2011). The cross-informant concordance and concurrent validity of the Borderline Personality Disorder Features Scale for Children in a community sample of boys. *Clinical Child Psychology and Psychiatry, 16*(3), 335-349.

Sharp, C., Pane, H., Ha, C., Venta, A., Patel, A., Sturek, J., & Fonagy, P. (2011). Theory of mind and emotion regulation difficulties in adolescents with borderline traits. *Journal of the American Academy for Child and Adolescent Psychiatry, 50*(6), 563-573.

Sharp, C., & Romero, C. (2007). Borderline personality disorder: A comparison between children and adults. *The Bulletin of the Menninger Clinic, 71*(2), 121-150.

Shiner, R. L. (2009). The development of personality disorders: Perspectives from normal personality development in childhood and adolescence. *Development and Psychopathology, 21,* 715-734.

Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research, 25,* 173-180.

Stevens, J. (1992). *Applied Multivariate Statistics for the Social Sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using Multivariate Statistics* (4th edition). Boston: Allyn & Bacon.

Tucker, L. R. & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika, 38*(1), 1-10.

Underwood, M. K., Beron, K. J., & Rosen, L. H. (2011). Joint trajectories for social and physical aggression as predictors of adolescent maladjustment: Internalizing symptoms, rule-breaking behaviors, and borderline and narcissistic personality features. *Development and Psychopathology, 23*(2), 659-678.

Wells, C., Adhyaru, J., Cannon, J., Lamond, M., & Baruch, G. (2010). Multisystemic therapy (MST) for youth offending, psychiatric disorder and substance abuse: Case examples from a UK MST team. *Child and Adolescent Mental Health, 15*(3), 142-149.

Werner, N. E., & Crick, N. R. (1999). Relational aggression and social-psychological adjustment in a college sample. *Journal of Abnormal Psychology, 108,* 615–623.

Westen, D., & Shedler, J. (1999a). Revising and assessing Axis II, part 1: 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 2: Toward an empirically based and clinically useful classification of personality disorders. *American Journal of Psychiatry, 156,* 273-285.

Westen, D., Shedler, J., Bradley, B., & DeFife, J.A. (2012). An empirically derived taxonomy for personality diagnosis: Bridging science and practice in conceptualizing personality. *American Journal of Psychiatry, 169(3),* 273-284.

Westen, D., Shedler, J., Durrett, C., Glass, S., & Martens, A. (2003). Personality diagnoses in adolescence: DSM-IV Axis II diagnoses and an empirically derived alternative. *American Journal of Psychiatry*, *160*, 952-966.

Wickline, V.B., Nowicki, S., Bollini, A. M., & Walker, E. F. (2012). Vocal and facial emotion decoding difficulties relating to social and thought problems: Highlighting schizotypal personality disorder. *Journal of Nonverbal Behavior, 36*(1), 59-77.

Widiger, T. A. (2010). Personality, interpersonal circumplex, and DSM–5: A commentary on five studies, *Journal of Personality Assessment, 92*(6), 528-532.

Wiznitzer, M., Verhulst, F.C., Van den Brink, W., Koeter, M., Van der Ende, J., Gie, l.R., & Koot, H.M. (1992). Detecting psychopathology in young adults. A comparison of the Young Adult Self-Report, the General Health Questionnaire and the Symptom Checklist as screening instruments. *Acta Psychiatrica Scandinavica, 86*, 32-37.

World Health Organization. (2010). *International Statistical Classification of Diseases and Related Health Problems 10th Revision.* Retrieved February 26, 2010.

Yu, C.Y., & Muthén, B. (2002). *Evaluation of model fit indices for latent variable models with categorical and continuous outcomes*. Technical report.

Zanarini, M. C. (2003). *Childhood Interview for DSM-IV Borderline Personality Disorder (CI-BPD).* Belmont, MA: McLean Hospital.

Zaslow, M., Halle, T., Martin, L., Cabrera, N., Calkins, J., Pitzer, L., et al. (2006). Child outcome measures in the study of child care quality. *Evaluation Review, 30,* 577-610.

Zelkowitz, P., Paris, J., Guzder, J., Feldman, R., Roy, C., & Rosval, L. (2007). A five-year follow-up of patients with borderline pathology of childhood. *Journal of Personality Disorders, 21*(6), 664-674.

Part 3: Critical Appraisal

Introduction

This paper extends the literature review and the empirical study presented in this volume by reflecting on key conceptual, clinical and methodological aspects of the research undertaken. Firstly, the personality disorder (PD) concept will be considered and the epistemological basis of this research identified. Next, developmental psychopathology will be briefly outlined as the conceptual framework underlying the studies presented in parts 1 and 2. This paper will then discuss the controversy surrounding the use of the PD diagnosis for young people, before reflecting on the proposed changes in the conceptualisation of PD in DSM-V (APA, 2012). Finally, I will discuss methodological aspects of this work, and in particular the use of a large dataset with information collected for routine outcome monitoring. Throughout the paper, personal experiences and reflections on the research process will be shared.

A note on terminology

Firstly, I would like to make some comments about the term “personality disorder” which I used throughout the thesis. In common with schizophrenia (Boyle, 1990), PD has been one of the most disputed terms in mental health. Although I am aware that the validity and reliability of the PD concept are now well established (Oldham, 2005) and that the term is widely accepted in academic clinical psychology and psychiatry, I still feel conflicted about its use, especially in the clinical context. My concern is that because personality makes us who we are, the term "personality disorder" may imply to the layman that the whole person with PD is flawed, which may be understandably experienced as a derogatory moral judgment (Appleby, 1988). When I asked myself how I would feel if I was given this diagnosis without having heard of the term before, my guess was that I would

probably feel insulted, confused, and rather hopeless about recovery. I can imagine how this rather contradictory, if not “disordered” term (Pilgrim, 2001) seems to have at least partly contributed, until recently, to the belief that PD is a condition for which no effective treatment could be offered. As a result, I found myself wondering whether a less controversial term could be used to describe these problems.

Alternative terms have been proposed for Borderline PD, such as Emotional Regulation Disorder, Emotional Intensity Disorder, and Emotion-Impulse Regulation Disorder. However, I found it hard to think of another term that would be descriptive of the wider range of personality problems. In what felt like a creative moment, I thought of the acronym “PIRSO” (Problems of Insecure Relationship to Self and Others), which may sound less definitive or deprecating. Nevertheless, I appreciate it is “a mouthful”, would be confusing to clients and open to interpretation. In any case, I am aware that the field is not in search of a new name, so I will use the term PD in this paper, carrying however the unanswered question of whether it could be replaced by a more appropriate and acceptable term. If and when the field identifies the need for a new name, I feel that consulting service users would be a good place to start.

Epistemological assumptions

Both the questions and the methods of the two papers are aligned with the positivist paradigm and reflect a nomothetic, probabilistic approach which is the dominant approach in research. The positivist position implies that a person has a definable, discoverable nature, and that conventional knowledge is based on objective unbiased observations of the world (Burr, 2003). However, a social constructionist perspective warns against the potential for this approach to trap individuals inside definitions of themselves which are limiting, or at worst expose them to oppressive practice (Burr, 2003).

The quantitative, cross-sectional, questionnaire-based research is the method mostly used in mainstream psychology, and the studies presented here are no exception.

Obviously, quantification is part of the scientific method and provides data that can be used to establish patterns, investigate sources of variation, and conduct statistical analyses to test hypotheses about the phenomena under investigation (Arnett, 2005). However, as the methods used in this research were exclusively quantitative, and given the large sample size, I was concerned that the object of study was reduced to numbers and that the humanity of the study's participants was somewhat lost by the end of the process.

As Arnett (2005) notices, this loss is perhaps especially unfortunate in the study of young people. On reflection, I focused on adolescence and emerging adulthood for my research and specialist clinical placement because I felt that it was an intriguing age of unique opportunities and vulnerabilities, and that the young are exciting and lively people who, given appropriate support, have a lot to offer. However, I felt that young people's voices and rich personalities were missing from both the papers studied to inform my research, and the papers I eventually produced. Obviously, as the focus of my studies was psychometric, quantitative methods were both suitable and necessary to enable generalisation. Nevertheless, inevitably, they do not reflect the distinctiveness of individual experience.

Identity development and an advanced capacity for self-reflection are among the hallmarks of adolescence (Arnett, 2005; Erikson, 1968), therefore what adolescents themselves say about their own experiences and personality is particularly important. However, I realise that at the end of this research I was left not knowing how for instance young people experienced completing the Achenbach System of Empirically Based Assessment (ASEBA) forms, what they thought about the questions they were asked, whether they felt that other questions could have been more pertinent to their developing personality, and what they understood by the terms "emerging PD", "personality pathology" or "PD traits".

Looking back at my PhD research, I noticed that participants' "voices" were missing from my studies at that time, too. However, having had the opportunity to work closely with young people in my clinical role, I do feel more in touch with young people's experiences now, even if not through research. In my future research, I wish to try harder to ensure that adolescents' voices are given prominence.

Another element that has not been attended to in this volume is the sociocultural element. This research adopted a largely monocultural perspective to adolescent development and mental health. The great majority of studies examined in this volume were carried out in North America or Europe, and the extent to which the concept of PD can be applicable to societies and cultural contexts outside the Western world has yet to be established (Millon & Grossman, 2005). In addition, it is worth keeping in mind that adolescent development can only be understood in the context of culture (Arnett, 2005). On the other hand, as far as the ASEBA is concerned, its forms have been translated into over 80 languages and extensive research has supported the ASEBA's applicability to research undertaken in diverse cultural contexts (Achenbach et al., 2008).

Moreover, the focus of this research was on psychopathology and risk, rather than protective factors and resilience. At a practical level, this was necessary as the adaptive functioning and social competence items of the ASEBA were not included in the papers reviewed and were not available for data analysis in the empirical study. Consequently, the focus was on the problem items and scales that may indicate personality pathology. However, because "Nothing is so fascinating or complicated as a trajectory of a human life" (p.550), as Hauser and Allen (2000) astutely noted, a more comprehensive approach to the complexity of adolescent development and psychopathology needs to include not only maladaptive, but also adaptive developmental processes and resilience mechanisms (Rutter, 1993). This is one of the key concepts in the field of developmental psychopathology that is briefly outlined below.

Developmental psychopathology as a common theme

The theoretical perspective underlying the two papers is in accordance with the field of developmental psychopathology; namely the study of mechanisms that cause developmental pathways to diverge toward pathological or typical outcomes (Cicchetti & Cohen, 1995). According to the developmental psychopathology perspective, bio-psycho-social factors in the individual, family, and wider social and ecological systems interact at different stages of development to bring about various outcomes (Cicchetti & Cohen, 1995). There is not only the gene-environment interaction, but also the developmental history of the individual that cumulatively influence the unfolding of future development (Sroufe, 2007). Another important feature of this perspective is that all psychopathology can be understood as a disturbance of normal functioning. Therefore, in order to understand developmental trajectories that result in pathological outcomes, e.g. PD diagnoses in adulthood, we must also understand and study those trajectories with alternative outcomes (Cicchetti & Cohen, 1995).

Related to this is the concept of *equifinality*, which implies that there are multiple pathways to similar manifest outcomes, and the concept of *multifinality*, which implies that the same risk and protective factors may lead to a number of different outcomes (Cicchetti & Rogosch, 1996). In the context of the current research, this means that a variety of developmental progressions, rather than a singular primary pathway, may eventuate in PD. For some young people, personality pathology may be situational but for some others it will be long-standing. What the developmental course is like in these young people who have such PD features, and whether their disturbance may be more transitory or more lifelong are questions for longitudinal research designs. Consequently, regardless of how reliable and valid a juvenile PD traits measure may be, it will never predict with certainty later outcome, but could be used to identify pathways probabilistically leading to PD in longitudinal research. Moreover, longitudinal designs are necessary for differentiating

between personal effects on the environment and environmental effects on the person (Rutter & Sroufe, 2000).

The issue of comorbidity raises the key question of what the co-occurrence of different forms of psychopathology means. Most PDs do not present to clinicians in a pure form and comorbidity is the rule rather than the exception (Bateman & Fonagy, 2008). Consequently, there remain important questions about diagnostic prediction from childhood and adolescence to adulthood (Copeland, Shanahan, Costello, & Angold, 2009). In accordance with the concept of homotypic prediction, PD traits in youth may predict PD over time, but heterotypic prediction is also possible, for instance PD traits in youth may predict other forms of later psychopathology, such as mood disorder or psychosis. Disorders with high co-occurrence, such as depression and PD, may derive from the same set of inter-correlated risk factors, but this may also mean that the presence of one form of psychopathology may, through its effects, constitute a risk mechanism for another form of psychopathology (Rutter & Sroufe, 2000). Prospective studies are again required to track individuals and delineate the unfolding of comorbidity over time.

Personality disorders in youth: conceptual and clinical dilemmas

Over the past three decades these ideas have taken a stronghold in the field of clinical psychology (Wright, 2009), and clinicians and clinical investigators increasingly recognise the need for a developmental life span perspective on disrupted personality functioning that goes beyond the atheoretical categorical system of the DSM (Blatt & Luyten, 2009).

Nonetheless, juvenile PD is a controversial diagnosis and a topic of heated debate. On the one hand, it has been argued that personality lacks cohesiveness and stability in children and adolescents (Miller, Muehlenkamp, & Jacobson, 2008), and that because the PD diagnosis was originally developed for adults, it does not take into account developmental issues associated with earlier stages of life (Shapiro, 1990). Another

objection regards the question of the durability of personality pathology in juveniles (Westen & Chang, 2000), and of course the possibility of labelling young people with these stigmatising diagnoses is the most serious clinical problem (Hinshaw, 2007).

People with PD have usually been viewed as “hard to help” or “difficult” (Cleary, Siegfried, & Walter, 2002; Kerr, 1999), and an awareness of a PD diagnosis has been associated with a clinician belief that people will be harder to manage (Newton-Howes, Weaver, & Tyrer, 2008). Predictably, a diagnosis of PD still carries great stigma, and those diagnosed can feel labelled by society as well as blamed by professionals (Haigh, 2002). PD has been considered “a very sticky label” by service users; once the diagnosis is recorded, it often remains indefinitely while at the same time professionals try to hide it (Haigh, 2002). Therefore, it is not surprising that clinicians feel very reluctant to diagnose PD in young people.

However, it is also possible that clinicians have gone one step too far. While on placement I noticed that mental health professionals working with young people are very reluctant to even contemplate the possibility of PD features. They seem to avoid mentioning the term in clinical meetings and case discussions, although in some cases young people’s clinical presentation was a very close match to PD prototypes. To a large extent, I too find myself uncomfortable with the notion of diagnosing PD in young people. At the same time, I have noticed how easy it is for these young people to develop a mental health “career” early in life and receive years of inappropriate treatment.

It has been argued that while some young people seem to move in and out of the PD diagnosis, a group of them that have stable characteristics of PD also appears to exist (Miller et al., 2008). If PDs are present in some juveniles, then we need to develop systematic knowledge of the aetiology and development of PD so that we can identify those at risk (Crick, Murray-Close, & Woods, 2005), and ascertain which interventions will be most efficacious (Cicchetti & Crick, 2009). In addition, if prospective longitudinal studies reveal precursors of emergent PDs, then preventive interventions need to be developed to

deflect vulnerable children from their pathological developmental trajectories. Furthermore, if stigma were to lessen, accurate and timely diagnosis may have benefits such as empowerment for the individual and family, a reduction in feelings of guilt and fostering of a sense of hope for intervention (Hinshaw, 2007).

Whilst it seems highly likely that earlier and more accurate detection of PD difficulties in young people could be beneficial, this should be weighed against the risks associated with stigmatisation and labelling. According to Cicchetti and Crick (2009), before assigning PD diagnoses to children, it may be more prudent to await the results of longitudinal investigations of the emergence of personality pathology across the life span. Should earlier diagnosis become more common, it would be necessary to adapt the diagnostic classifications to incorporate developmentally appropriate markers of PD (Chanen et al., 2004). Most importantly, PD in youth should be regarded as an emerging style of relating to others that is problematic, but at the same time amenable to change and intervention planning with tremendous psychosocial benefits (Burnette, South, & Reppucci, 2007).

I feel that clinicians of my generation have strong reasons to be hopeful about their work with people with PD. Existing evidence-based interventions include Mentalization-Based Therapy (Bateman & Fonagy, 2004) and Dialectical Behavioural Therapy (Linehan, 1993). I feel particularly fortunate to have studied in a university where the former therapy was developed, and to have worked with mentors that revolutionised the field. On the other hand, I am mindful of the “loyalty” I feel towards a theoretical perspective that emphasises the developmental origins of PD. At times I wonder whether I am able to fully appreciate less dominant, alternative discourses which criticise the PD concept and the medical perspective, whilst emphasising the role of social disadvantage and oppressive practice in causing and maintaining mental health problems.

The future of personality disorders

Whilst planning and carrying out this research, work toward a fifth edition of the DSM was well underway. The research agenda prepared for DSM-V emphasises the need for a better understanding of the developmental origins of PDs (Crawford et al., 2008), and the proposed changes to the conceptualisation of PDs also include substantial revisions to the current categorical system. The rationale for the proposed changes in PD assessment and diagnosis derives from increasing evidence favouring dimensional over categorical conceptualisations of PD (Clark, 2007). Moreover, the proposed changes stem from several problems with the existing DSM-IV-TR system, including "an unsubstantiated and nonspecific definition of and general criteria for PD; the lack of a PD specific, clinically useful, severity measure; excessive diagnostic comorbidity among DSM-IV-TR PDs; limited validity of some existing types; arbitrary diagnostic thresholds; within-disorder heterogeneity; inadequate coverage of the range of PD pathology, and instability of current diagnostic criteria sets" (Skodol et al., 2011, p.24).

The proposed reformulation holds that PDs "represent the failure to develop a sense of self-identity and the capacity for interpersonal functioning that are adaptive in the context of the individual's cultural norms and expectations". The new DSM-V assessment model for personality psychopathology identifies core impairments in personality functioning, pathological personality traits, and prominent pathological personality types.

In particular, it consists of the following parts: 1) five identified severity levels of personality functioning, based on degrees of impairment in core self and interpersonal capacities, ranging from normal to severely impaired, 2) prototype descriptions of six major personality (disorder) types, i.e. antisocial, avoidant, borderline, narcissistic, obsessive-compulsive and schizotypal 3) a personality trait assessment including five broad, higher order personality trait domains (i.e. negative affectivity, detachment, antagonism, disinhibition vs. compulsivity and psychoticism), with 3-9 lower order, more specific trait facets within each domain, for a total of 29 specific trait facets, 4) generic criteria for PD

consisting of severe deficits in self-differentiation and integration and in the capacity for interpersonal relatedness, and 5) measures of adaptive functioning.

The proposed revisions have been designed “for flexible use to maximise clinical utility”, expecting that “even a busy clinician with limited time or expertise in the assessment of personality or PDs should be able to decide whether a personality-related problem exists and how severe it is” (Skodol et al., 2011, p.24). However, objections have been raised against the new system by numerous experts in the field (e.g. Shedler et al., 2010; Widiger, 2011). Concerns are primarily in regard to the omission of several DSM-IV TR (APA, 2000) PD diagnoses from the manual, the abandonment of diagnostic criterion sets, and the inclusion of a “needlessly” complicated trait-based dimensional model that lacks empirical support and adequate clinical rationale and utility (Shedler et al., 2010).

As a trainee clinician and researcher, I have found this controversy intriguing. The fact that people who have dedicated their careers, if not lives, to the understanding and treatment of PDs may have so opposing views about the appropriateness and usefulness of the changing diagnostic system has made me realise the complexity of the phenomena and decisions in question. As Clark (2007) remarked, personality - both adaptive and maladaptive - is too complex to be assessed comprehensively from a single perspective, so I appreciate that difference and disagreement should not be regarded negatively but welcomed. I have also come to appreciate that flaws are inevitable in any diagnostic system, and that a diagnostic manual may not directly translate into clinical practice. One would hope, however, that compared to the previous diagnostic system, the revised manual should at least facilitate clinical judgment, increase clinical utility and most importantly improve client care (First et al., 2004).

In clinical practice, PDs are still not being formally recognised (Woodrow, Shinner, & Tai, 2008), and this is in keeping with my experience. During training, I worked with at least five clients who had strong PD traits and whose life seemed to be significantly impaired as a result. However, none of them had received a PD diagnosis, but they had

instead a collection of other diagnoses, for which they were receiving (often inappropriate) treatment. Hopefully, our assessment practices will become more effective as we understand PD better and develop suitable tools to assist clinicians, given the biases pertinent to clinical assessment (Dutra, Campbell, & Westen, 2004). At the same time, I have noticed that apart from psychologists, clinicians from other backgrounds are quite reluctant to use psychometric tools in assessment. Educating our colleagues from other professions about these tools, and supervising their use accordingly is a role for clinical psychologists to champion.

Using second-hand data: limitations and benefits

My empirical study was conducted in a setting that values the use of psychometric tools (e.g. ASEBA) for both assessment and outcome monitoring purposes. The data analysed and presented in the second paper were primarily collected for the purposes of routine outcome monitoring. This means that data collection was not designed to answer the questions of this thesis. On the one hand, this is a significant limitation as the operationalisation of a key variable, namely PD, was insufficient. Furthermore, I had no influence on the sampling process, and at times felt unsure about the extent to which I "owned" the research produced.

On the other hand, I had the privilege of accessing a really large database that took about twenty years of systematic work and the effort of numerous clinicians and administration staff to put together. In addition, rearranging and combining the databases took a significant amount of time, which would have been very difficult to invest if research time had to be allocated on collecting new data. Furthermore, the opportunity to collaborate with the Brandon Centre over the years gave me insights into the importance of routine outcome monitoring in the voluntary sector. I have come to understand that outcome data not only determine a service's treatment effectiveness but they also help to secure funding for existing and new services and projects (Baruch & Vrouva, 2010).

Moreover, given the gap between academic research and the reality of everyday clinical practice, I felt that this project benefited from (and in return may contribute to) both worlds.

Concluding remarks

On a personal note, conducting the studies presented in this volume alongside my clinical training reinforced my pre-existing interest in research, and at the same time made me more aware of the different challenges and rewards inherent in research and clinical work.

Beyond their differences, research and clinical work both require commitment and flexibility in the face of uncertainty and frequent disappointment. For instance, whilst I had hoped that this research would lead to the development of a psychometrically robust PD traits ASEBA scale, I had to remind myself that although this aim was not achieved, the research produced could still be of value.

Research ultimately raises more questions than it answers, and this can be the case with clinical work, too. Clinical psychologists' training as scientists-practitioners enables them to face such challenges, as long as we remember that categorical descriptions and assessment tools are only useful when they aid formulation and guide individualised psychosocial intervention that respects and celebrates personality's rich complexity.

References

Achenbach, T. M., Becker, A., Döpfner, M., Heiervang, E., Roessner, V., Steinhausen, H.C., & Rothenberger, A. (2008). Multicultural assessment of child and adolescent psychopathology with ASEBA and SDQ instruments: research findings, applications, and future directions. *Journal of Child Psychology and Psychiatry, 49*(3), 251-275.

American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders* (*4th edition, text revision*). Washington, DC: American Psychiatric Association.

American Psychiatric Association (2012). Personality and Personality Disorders. http://www.dsm5.org/PROPOSEDREVISIONS/Pages/PersonalityandPersonalityDisorders.aspx

Appleby, L. G. (1988). Personality disorder: The patients psychiatrists dislike. *British Journal of Psychiatry, 153*, 44-49.

Arnett, J. J. (2005). The Vitality Criterion: A New Standard of Publication for Journal of Adolescent Research *Journal of Adolescent Research*, *20*(3), 3-7.

Baruch, G., & Vrouva, I. (2010). Collecting routine outcome data in a psychotherapy community clinic for young people: Findings from an ongoing study. *Child and Adolescent Mental Health, 15*(1), 30-36.

Bateman, A. W., & Fonagy, P. (2004). *Psychotherapy for borderline personality disorder: Mentalization based treatment*. Oxford: Oxford University.

Bateman A. W., & Fonagy, P. (2008). Comorbid Antisocial and Borderline Personality Disorders: Mentalization-Based Treatment. *Journal of Clinical Psychology, 64*(2),181-194.

Blatt S. J., & Luyten, P. (2009). A structural-developmental psychodynamic approach to psychopathology: Two polarities of experience across the life span *Development and Psychopathology, 21*(3), 793-814.

Boyle, M. (1990). *Schizophrenia: A Scientific Delusion? (*London, Routledge).

Burnette, M. L., South, S. C., & Reppucci, N. D. (2007). Cluster B personality pathology in incarcerated girls: Structure, comorbidity, and aggression. *Journal of Personality Disorders 21*(3), 262-272.

Burr, V. (2003). What is social constructionism? In V. Burr (Ed). *Social Constructionism 2nd Edition*. (pp 1-27). Psychology Press.

Chanen, A. M., Jackson, H. J., McGorry, P. D., Allot, K. A., Clarkson, V., & Yuen, H. P. (2004). Two-year stability of personality disorder in older adolescent outpatients. *Journal of Personality Disorders, 18*(6), 526-541.

Cicchetti, D., & Cohen, D. J. (1995). Perspectives on developmental psychopathology. In D. Cicchetti, & D. J. Cohen (Eds.), *Developmental psychopathology: Vol. 1. Theory and methods* (pp. 3-17). New York: Wiley.

Cicchetti, D., & Crick, N. R. (2009). Editorial. Precursors and diverse pathways to personality disorder in children and adolescents. *Development and Psychopathology, 21*, 683-685.

Cicchetti, D., & Rogosch, F. A. (1996). Equifinality and multifinality in developmental psychopathology. *Development and Psychopathology, 8*(4), 597-600.

Clark, L.A. (2007). Assessment and diagnosis of personality disorder: perennial issues and an emerging reconceptualization. *Annual Review of Psychology, 58,* 227-57.

Cleary, M., Siegfried, N., & Walter, G. (2002). Experience, knowledge and attitudes of mental health staff regarding clients with a borderline personality disorder. *International Journal of Mental Health Nursing, 11*(3), 186-191.

Copeland, W. E., Shanahan, L, Costello, J. E., & Angold, A. (2009). Which Childhood and Adolescent Psychiatric Disorders predict which Young Adult Disorders? *Archives of General Psychiatry, 66*(7), 764-772.

Crawford, T.N., Cohen, P., First, M.B., Skodol, A.E., Johnson, J.G., & Kasen, S. (2008). Comorbid Axis I and Axis II disorders in early adolescence. *Archives of General Psychiatry, 65*(6), 641-648.

Crick, N. R., Murray-Close, D., & Woods, K. (2005). Borderline personality features in childhood: A short-term longitudinal study. *Development and Psychopathology, 17*(4), 1051-1070.

Dutra, L., Campbell, L., & Westen, D. (2004). Quantifying Clinical Judgment in the Assessment of Adolescent Psychopathology: Reliability, Validity, and Factor Structure of the Child Behavior Checklist for Clinician Report. *Journal of Clinical Psychology, 60*(1),

65-85.

Erikson, E. H. (1968). *Identity: Youth and crisis.* New York: W. W. Norton.

First, M.B., Pincus, H.A., Levine, J.B., Williams, J.B.W, Ustun, B., & Peele, R. (2004). Clinical utility as a criterion for revising psychiatric diagnoses. *American Journal of Psychiatry*, *161*, 946-954.

Haigh, R. (2002). *Services for people with Personality disorder: The thoughts of service users.* Key text on NIMHE website.

Hauser, S. T., & Allen, J. P. (2000). Overcoming Adversity in Adolescence: Narratives of Resilience. *Psychoanalytic Inquiry, 26*, 549-576.

Hinshaw, S.P. (2007). Evidence from Scientific Investigations. In S.P. Hinshaw (Ed.), *The Mark of Shame: Stigma of Mental Illness and an Agenda for Change.* (pp 114). Oxford University Press.

Kerr, I. B. (1999). Cognitive analytic therapy for borderline personality disorder in the context of a Community Mental Health Team: individual and organisational psychodynamic implications. *British Journal of Psychotherapy*, *15*, 425–438.

Linehan, M. M. (1993). *Cognitive-behavioral treatment of borderline personality disorder.* New York: Guilford.

Miller, A. L., Muehlenkamp, J. J., & Jacobson, C. M. (2008). Fact or fiction: Diagnosing borderline personality disorder in adolescents. *Clinical Psychology Review, 28,* 969-981.

Millon, T. & Grossman, S. (2005). Sociocultural factors. In J. M. Oldham, A. E. Skodol & D. S. Bender (Eds.), *The American Psychiatric Publishing textbook of personality disorders* (pp. 223-235). Washington, DC: American Psychiatric Press.

Newton-Howes, G., Weaver, T., & Tyrer, P. (2008). Attitudes of staff towards patients with personality disorder in community mental health teams. *Australian and New Zealand Journal of Psychiatry, 42,* 572-577.

Oldham, J. M. (2005). Personality Disorders. Recent History and Future Directions. In J. M. Oldham, A. E. Skodol & D. S. Bender (Eds.), *The American Psychiatric Publishing textbook of personality disorders* (pp. 3-15). Washington, DC: American Psychiatric Press.

Pilgrim, D. (2001). Disordered personalities and disordered concepts. *Journal of Mental Health, 10*(3), 253-265.

Rutter, M. (1993). Resilience: Some conceptual considerations. *Journal of Adolescent Health, 14,* 626-631.

Rutter, M., & Sroufe, L. A. (2000). Developmental psychopathology: Concepts and challenges. *Development and Psychopathology, 12,* 265-296.

Shapiro, T. (1990). Debate forum-resolved: Borderline personality disorder exists in children under twelve. *Journal of the American Academy of Child and Adolescent Psychiatry*, *29*, 478−483.

Shedler, J., Beck, A., Fonagy, P., Gabbard, G. O., Gunderson, J., Kernberg, O., et al. (2010). Personality Disorders in DSM-5. *American Journal of Psychiatry, 167*(9), 1026-1028.

Skodol, A. E., Bender, D.S., Oldham, J. M., Clark, L. E., Morey, L. A., Verheul, R., et al. (2011). Proposed Changes in Personality and Personality Disorder Assessment and Diagnosis for *DSM-5* Part II: Clinical Application. *Personality Disorders: Theory, Research, and Treatment, 2*, 23-40.

Sroufe, L. A. (2007). The place of development in developmental psychopathology. In A.S. Masten (Ed.), *Multilevel dynamics in developmental psychopathology: The Minnesota Symposia on Child Psychology* (vol. 34, pp. 285-299). Mahwah, NJ: Lawrence Erlbaum Associates.

Westen, D., & Chang, C. (2000). Personality pathology in adolescence: A review. *Adolescent Psychiatry, 25,* 61-100.

Widiger, T. A. (2011). A Shaky Future for Personality Disorders. *Personality Disorders: Theory, Research, and Treatment, (2)*1, 54–67.

Woodrow, C., Shinner, M., & Tai, S. (2008). Care pathway for people with a diagnosis of personality disorder. *Clinical Psychology Forum, 191*, 39-43.

Wright, C. (2009). Where does it all go wrong? *A Developmental Perspective on Personality Difficulties*. Unpublished Doctoral Dissertation, University College London, UK.

APPENDIX

Appendix I

Letter of Ethical Approval

UCL RESEARCH ETHICS COMMITTEE
GRADUATE SCHOOL OFFICE

UCL

Professor Peter Fonagy
Research Department of Clinical, Educational and Health Psychology
UCL

14 July 2011

Dear Professor Fonagy

Notification of Ethical Approval
Ethics Application: 0596/002: Predictors of emerging personality disorder (PD) and the utility of the Achenbach System of Empirically Based Assessment (ASEBA) in the study and diagnosis of emerging PD in a community-based counselling and psychotherapy service for adolescents and young adults

I am pleased to confirm that in my capacity as Chair of the UCL Research Ethics Committee I have approved your project for the duration of the study (i.e. until September 2012).

However, approval is subject to the following conditions:

1. You must seek Chair's approval for proposed amendments to the research for which this approval has been given. Ethical approval is specific to this project and must not be treated as applicable to research of a similar nature. Each research project is reviewed separately and if there are significant changes to the research protocol you should seek confirmation of continued ethical approval by completing the 'Amendment Approval Request Form'.

The form identified above can be accessed by logging on to the ethics website homepage: http://www.grad.ucl.ac.uk/ethics/ and clicking on the button marked 'Key Responsibilities of the Researcher Following Approval'.

2. It is your responsibility to report to the Committee any unanticipated problems or adverse events involving risks to participants or others. Both non-serious and serious adverse events must be reported.

   Reporting Non-Serious Adverse Events
   For non-serious adverse events you will need to inform Helen Dougal, Ethics Committee Administrator (ethics@ucl.ac.uk), within ten days of an adverse incident occurring and provide a full written report that should include any amendments to the participant information sheet and study protocol. The Chair or Vice-Chair of the Ethics Committee will confirm that the incident is non-serious and report to the Committee at the next meeting. The final view of the Committee will be communicated to you.

   Reporting Serious Adverse Events
   The Ethics Committee should be notified of all serious adverse events via the Ethics Committee Administrator immediately the incident occurs. Where the adverse incident is unexpected and serious, the Chair or Vice-Chair will decide whether the study should be terminated pending the opinion of an independent expert. The adverse event will be considered at the next Committee meeting and a decision will be made on the need to change the information leaflet and/or study protocol.

On completion of the research you must submit a brief report (a maximum of two sides of A4) of your findings/concluding comments to the Committee, which includes in particular issues relating to the ethical implications of the research.

With best wishes for the research.

Yours sincerely

Sir John Birch
Chair of the UCL Research Ethics Committee

Cc: Ioanna Vrouva

UCL Research Ethics Committee, c/o The Graduate School, North Cloisters, Wilkins Building
University College London Gower Street London WC1E 6BT
Tel: +44 (0)20 7679 7844 Fax: +44 (0)20 7679 7043
ethics@ucl.ac.uk
www.ucl.ac.uk/gradschool

APPENDIX II

Table 1. *Hierarchical logistic regression predicting PD traits vs. non-PD traits (B>0 indicates that the variable predicts PD traits) (self-reports)*

							95% C.I. for Odds Ratio	
		B	SE of *B*	Wald Test	*p* value	Odds Ratio	Lower	Upper
Step 1	Age	.107	.022	23.623	.000	1.113	1.066	1.162
	Female gender	.650	.136	22.939	.000	1.915	1.468	2.498
	Neurotic, stress-related or somatoform disorder	1.130	.184	37.775	.000	3.096	2.159	4.439
	Mood disorder	.929	.152	37.577	.000	2.532	1.881	3.407
	Somatic Problems	-.005	.008	.393	.531	.995	.980	1.011
Step 2	Anxious/Depressed	.009	.008	1.374	.241	1.009	.994	1.025
	Thought Problems	.013	.008	3.106	.078	1.013	.999	1.028
	Attention Problems	.000	.009	.000	.988	1.000	.983	1.017
	Delinquent Behaviour	.032	.009	12.713	.000	1.033	1.015	1.051
	Aggressive Behaviour	-.005	.008	.318	.573	.995	.979	1.012
	Withdrawn	.011	.008	1.931	.165	1.011	.995	1.028
	Constant	-7.930	.804	97.312	.000	.000		

Table 2. *Hierarchical logistic regression predicting PD (using expert defined criteria) vs. non-PD (B>0 indicates that the variable predicts PD) (self-reports)*

						95% C.I. for Odds Ratio	
		B	SE of *B*	Wald Test	*p* value	Odds Ratio	Lower	Upper
Step 1	Age	.107	.052	4.207	.040	1.113	1.005	1.234
	Mood disorder	.522	.304	2.949	.086	1.686	.929	3.060
Step 2	Somatic Problems	.030	.017	3.194	.074	1.031	.997	1.065
	Anxious/Depressed	-.037	.018	4.180	.041	.964	.930	.998
	Thought Problems	.051	.015	11.208	.001	1.053	1.021	1.085
	Attention Problems	.001	.020	.002	.965	1.001	.963	1.040
	Delinquent Behaviour	.078	.020	15.953	.000	1.082	1.041	1.124
	Aggressive Behaviour	.031	.018	3.179	.075	1.032	.997	1.068
	Withdrawn	-.020	.019	1.132	.287	.980	.944	1.017
	Constant	-13.920	1.814	58.860	.000	.000		

Table 3. *Hierarchical logistic regression predicting PD traits vs. non-PD traits (B>0 indicates that the variable predicts PD traits) (therapist reports)*

							95% C.I. for Odds Ratio	
		B	SE of *B*	Wald Test	*p* value	Odds Ratio	Lower	Upper
Step 1	Age	.121	.026	21.486	.000	1.128	1.072	1.187
	Female gender	.661	.159	17.361	.000	1.937	1.419	2.643
	Neurotic, stress-related or somatoform disorder	1.094	.216	25.619	.000	2.986	1.955	4.560
	Mood disorder	.747	.177	17.827	.000	2.110	1.492	2.985
Step 2	Withdrawn	.007	.011	.390	.532	1.007	.986	1.028
	Somatic Problems	-.018	.009	4.088	.043	.982	.965	.999
	Anxious/Depressed	.006	.010	.391	.532	1.006	.987	1.026
	Thought Problems	.061	.011	33.769	.000	1.063	1.041	1.085
	Attention Problems	-.012	.013	.801	.371	.989	.964	1.014
	Delinquent Behaviour	.034	.012	8.103	.004	1.035	1.011	1.060
	Aggressive Behaviour	.050	.013	15.724	.000	1.051	1.026	1.078
	Constant	-12.948	1.183	119.768	.000	.000		

Table 4. *Hierarchical logistic regression predicting PD (using expert defined criteria) vs. non-PD (B>0 indicates that the variable predicts PD) (therapist reports)*

							95% C.I. for Odds Ratio	
		B	SE of *B*	Wald Test	*p* value	Odds Ratio	Lower	Upper
Step 1	Age	.076	.053	2.055	.152	1.079	.973	1.197
	Mood disorder	.142	.318	.199	.655	1.153	.618	2.152
Step 2	Withdrawn	-.018	.023	.663	.415	.982	.939	1.026
	Somatic Problems	-.009	.018	.221	.638	.991	.956	1.028
	Anxious/Depressed	-.012	.021	.327	.567	.988	.948	1.030
	Thought Problems	.031	.021	2.224	.136	1.031	.990	1.074
	Attention Problems	.015	.023	.431	.512	1.015	.971	1.062
	Delinquent Behaviour	.078	.023	11.871	.001	1.082	1.034	1.131
	Aggressive Behaviour	.071	.022	10.770	.001	1.073	1.029	1.119
	Constant	-14.654	2.214	43.812	.000	.000		

Table 5. *Hierarchical logistic regression predicting PD traits vs. non-PD traits (B>0 indicates that the variable predicts PD traits) (significant other reports)*

							95% C.I. for Odds Ratio	
		B	SE of *B*	Wald Test	*p* value	Odds Ratio	Lower	Upper
Step 1	Age	.115	.027	18.175	.000	1.122	1.064	1.183
	Female gender	.648	.164	15.663	.000	1.911	1.387	2.634
	Neurotic, stress-related or somatoform disorder	1.083	.218	24.587	.000	2.953	1.925	4.529
	Mood disorder	.912	.185	24.340	.000	2.488	1.732	3.574
Step 2	Withdrawn	.022	.011	4.321	.038	1.022	1.001	1.044
	Somatic Problems	-.012	.008	1.964	.161	.988	.972	1.005
	Anxious/Depressed	.003	.010	.099	.753	1.003	.984	1.023
	Thought Problems	.027	.010	7.842	.005	1.028	1.008	1.048
	Attention Problems	.006	.012	.253	.615	1.006	.982	1.031
	Delinquent Behaviour	.042	.012	11.717	.001	1.043	1.018	1.068
	Aggressive Behaviour	-.019	.013	2.160	.142	.981	.957	1.006
	Constant	-8.880	1.089	66.481	.000	.000		

Table 6. *Hierarchical logistic regression predicting PD (using expert defined criteria) vs. non-PD (B>0 indicates that the variable predicts PD) (significant other reports)*

							95% C.I. for Odds Ratio	
		B	SE of *B*	Wald Test	*p* value	Odds Ratio	Lower	Upper
Step 1	Age	.129	.070	3.364	.067	1.138	.991	1.306
	Mood disorder	-.087	.378	.053	.818	.917	.437	1.924
Step 2	Withdrawn	.017	.023	.566	.452	1.017	.973	1.064
	Somatic Problems	-.006	.019	.099	.753	.994	.957	1.032
	Anxious/Depressed	-.025	.023	1.242	.265	.975	.933	1.019
	Thought Problems	.080	.024	10.900	.001	1.084	1.033	1.137
	Attention Problems	.049	.023	4.501	.034	1.050	1.004	1.099
	Delinquent Behaviour	.026	.027	.926	.336	1.026	.974	1.081
	Aggressive Behaviour	.009	.025	.123	.726	1.009	.961	1.059
	Constant	-15.402	2.705	32.419	.000	.000		