Refinement of the children's social behavior questionnaire (CSBQ)

Hartman, CA; Luteijn, E; Serra, M; Minderaa, R; Luteijn, F.

Published in:
Journal of Autism and Developmental Disorders

DOI:
10.1007/s10803-005-0072-z

Journal of Autism and Developmental Disorders, Vol. 36, No. 3, April 2006 (© 2006)
DOI 10.1007/s10803-005-0072-z
Published Online: April 15, 2006

Refinement of the Children's Social Behavior Questionnaire (CSBQ): An Instrument that Describes the Diverse Problems Seen in Milder Forms of PDD

Catharina A. Hartman,[1,3] Ellen Luteijn,[2] Marike Serra,[1] and Ruud Minderaa[1]

The objective of this study was to refine the Children's Social Behavior Questionnaire (CSBQ), to reduce its length, and to verify its psychometric properties. The CSBQ is a questionnaire for parents or caregivers of children with PDD. The items describe a broad range of features that are typical of PDD, particularly in its milder forms. Based on conceptual judgment and factor analyses, the number of items was reduced from 96 to 49. Six subscales were constructed to allow a differentiated description of PDD problems. Estimates for internal, test–retest, and inter-rater reliability, and for convergent and divergent validity were good. Different clinical and control groups showed the hypothesized patterns in nature and degree of their problems.

KEY WORDS: PDD; PDDNOS; ADHD; mental retardation; CSBQ; psychometrics.

INTRODUCTION

The two major classification systems, the DSM-IV (APA, 1994; DSM-IV-TR, 2002) and the ICD-10 (WHO, 1994), have demonstrated improved sensitivity and specificity for Pervasive Developmental Disorders (PDDs) as compared to older systems (Volkmar, Klin, & Cohen, 1997). However, the present diagnostic criteria for specific PDDs (including Autistic Disorder, Asperger's Disorder, Childhood Disintegrative Disorder and Rett's Disorder) still exclude many individuals who have problems in social interaction and communication, or who have stereotyped and restricted patterns of behavior, yet do not fully meet the criteria for any of the above categories. These problems are currently classified as Pervasive Developmental Disorder Not Otherwise Specified (PDDNOS).

The DSM-IV-TR (2002) states that PDDNOS should be used "when there is a severe and pervasive impairment in the development of reciprocal social interaction associated with impairment in either verbal or non-verbal communication skills or with the presence of stereotyped behavior, interests, and activities, but the criteria are not met for a specific Pervasive Developmental Disorder, Schizophrenia, Schizotypical personality disorder, or Avoidant personality disorder." For example, this category includes "atypical autism"—a presentation that does not meet the criteria for Autistic disorder because of late age at onset, atypical symptomatology, subthreshold symptomatology, or all of these (APA, 1994). Subthreshold problems refer either to situations in which not enough symptoms are present to classify the problems as autism or to situations in

[1] Department of Psychiatry, Faculty of Medicine, University of Groningen, Groningen, The Netherlands.
[2] School for Special Education for Children and Adolescents with Language and Communication Problems, Arnhem, The Netherlands.
[3] Correspondence should be addressed to: Catharina A. Hartman, Department of Psychiatry, University Medical Center Groningen, P.O. Box 30.001, 9700 RB, Groningen, The Netherlands; e-mail: c.hartman@accare.nl

which the symptoms are present, but in a more subtle form. Thus, PDDNOS is a catchall diagnosis for those who do not meet the criteria for any specific PDD. It is mainly described in terms of what it is not, and there is no exact description of the problems, nor are there clear criteria with regard to the minimum number or severity of symptoms that must be present to qualify for the diagnosis (Walker *et al*., 2004).

Because of this lack of specific inclusion criteria, the PDDNOS classification is applied to individuals with a wide range of symptoms, and the reliability of the diagnosis is low (e.g., Beglinger & Tristram, 2001; Luteijn, Luteijn, Jackson, Volkmar, & Minderaa, 2000a, 2000b; Mahoney *et al*., 1998; Prior *et al*., 1998; Towbin, 1997). The lack of diagnostic consensus on PDDNOS, the heterogeneity of the clinical group, and the low reliability of the diagnosis combine to have a negative effect on research (Mayes, Volkmar, Hooks, & Cicchetti, 1993). A further diagnostic complication that hinders research is the frequent co-occurrence of PDDNOS with other disorders or maladaptive behaviors, such as mental retardation (Waterhouse *et al*., 1996), ADHD (Althaus, 2000, Perry, 1998), or language problems (Bishop & Norbury, 2002). This makes it difficult to distinguish between the symptoms that fall within the boundaries of PDDNOS and those that are primarily related to other conditions. Research on PDDNOS is very much warranted, however, if only because the PDD-NOS group includes substantially more children than does the population of those who meet the stringent diagnostic definition of autism (Chakrabarti & Fombonne, 2001; Volkmar *et al*., 1997).

In addition to these diagnostic problems, the relative lack of standardized instruments for charting the wide variety of symptoms seen in children with PDDNOS stands in the way of research. A number of instruments have been developed that are suitable for screening and diagnosing the more severe variants of PDD, including the Social Communication Questionnaire (SCQ; Berument, Rutter, Lord, Pickles, & Bailey, 1999; Rutter, Bailey, & Lord, 2003) and the ADI-R (Lord, Rutter, & LeCouteur, 1994). The major problem with using these instruments to describe the symptoms of children with PDDNOS is that the items in these instruments, directly derived from the DSM-IV or ICD-10 criteria for Autistic Disorder, are intended to screen for autism (as with the SCQ) or establish caseness (as with the ADI-R). While these instruments may be applicable to the subgroup of children with PDDNOS who have more severe problems, the items do not tap the more subtle problems of the less severe subgroup of PDDNOS. Further, the dichotomous scoring formats applied in these instruments do not allow the ordinal rating of the extent to which problem behaviors occur.

For the reasons presented above, the development of standardized measurement instruments to chart the various forms of social, communication, and stereotyped problem behaviors in children with milder forms of PDD constitutes a valuable contribution to research in this field. With this in mind, we developed the Children's Social Behavior Questionnaire (CSBQ) (Luteijn, Jackson, Volkmar, & Minderaa, 1998; Luteijn *et al*., 2000a). In the same period, Constantino and colleagues developed the Social Responsiveness Scale (SRS), which also seeks to provide a clinical characterization for the behavior of children whose deficits fall below the threshold for a full diagnosis of autism (Constantino, Przybeck, Friesen, & Todd, 2000).

In its original form, the CSBQ contained 96 items referring to a broad range of behavioral problems observed in children between the ages of 4 and 18 who have PDD. The items cover the entire spectrum of PDD problems, but the instrument emphasizes the milder and subtler variants seen in children with PDDNOS. The items are rated in an ordinal rather than a discrete (present/not present) fashion, in order to establish the extent to which problems are present. In 2000, the psychometric properties of the 96-item version of the CSBQ were reported. With few exceptions (see page 14), indices of reliability and validity of the instrument were good (Luteijn *et al*., 2000a). Nonetheless, there are several indications that further study and refinement of the CSBQ are necessary.

The first reason for refining the CSBQ is the current availability of data on a large number of children. The present paper uses factor analysis to derive the CSBQ subscales. Factor analysis of a large number of skewed, ordinally distributed items (a common situation in the measurement of child psychopathology) requires large samples in order to derive a factor structure that can be replicated in future samples (Hartman, 2000). A second reason is the instrument's length (i.e., 96 items), which weakens its applicability for both research and clinical purposes. A more concise instrument would therefore be advantageous for practical reasons. A third reason is that further refinement of the subscales would improve conceptual transparency. For example, in earlier studies, one of the subscales (''social insight problems'') contained items that refer to both problems of

social insight and inattentive behaviors, thereby impeding the interpretation of the scores on this subscale (i.e., high scores could indicate reduced social insight, attention problems, or both). Finally, refining the CSBQ could provide further insight into the extent to which different (clinical) groups (e.g., those diagnosed with PDDNOS, ADHD, internalizing disorders, high functioning autism, mental retardation, mental retardation with PDD, as well as typically developing children) differ in the nature and degree of PDD-related problems. The identification of distinctive score profiles for different clinical groups would attest to the construct validity of the instrument.

MATERIALS AND METHOD

Sample

The total sample (n = 3,407) consisted of three subsamples. Subsample 1 included 2,271 children who had various emotional, behavioral and developmental psychiatric problems, including a substantial number of children with PDD(NOS). Subsample 2 included 904 mentally retarded children (with and without PDD) and children functioning at a borderline intellectual level. Subsample 3 included 232 typically developing children.

Information from all of the 3,407 children was included in the factor analytic studies of the CSBQ. For 3,234 children both the CSBQ and the Child Behavior Checklist (CBCL; Achenbach, 1991; Verhulst, Van der Ende, & Koot, 1996) were completed. The information obtained from these children was used in the simultaneous factor analysis of CSBQ and CBCL items. In approximately 49% of the cases, the rater was the mother; in 5% of the cases, it was the father, and in 43% of the cases, both parents completed the questionnaires together. For the remaining 3%, the rater was a caregiver or a stepparent, with or without a biological parent. The CSBQ and the CBCL, as well as the factor analyses, are described in more detail below.

Only children whose psychiatric problems could be classified relatively unambiguously (i.e., they could be classified into a single Axis 1 DSM-IV diagnostic category) were included in comparisons across clinical groups. The two exceptions to this rule were (1) a group of children who had dual diagnoses of PDD-NOS and ADHD, and (2) a group of children who had dual diagnoses of PDD(NOS) and Mental Retardation. Because ADHD and Mental Retardation frequently occur along with PDD(NOS), these comorbid groups are relevant to the present research. As described below, we allowed for other comorbid problems in Subsample 1, but only to the extent that these problems were clearly subordinate to the main diagnosis. Following these rules, 1,605 (71%) of the children in Subsample 1 could be classified unambiguously, as could 733 (81%) of the children in Subsample 2. None of the children in Subsample 3 had experienced significant behavioral or emotional problems. The following sections provide more details about these subsamples and how they were recruited.

Subsample 1 ($N = 2{,}271$) included children diagnosed with high functioning autism (HFA) ($n = 102$), PDDNOS ($n = 544$), ADHD ($n = 586$), dual diagnoses of ADHD and PDD ($n = 164$), internalizing disorders (ID) ($n = 209$), and other psychiatric problems ($n = 666$).

Most of the children and adolescents in this subsample visited an outpatient clinic for child and adolescent psychiatry in Groningen, the Netherlands ($n = 2{,}174$; 96%). Most of the high functioning autistic subjects (HFA group), however, were recruited elsewhere ($n = 97$; 95%; described further below). Between June 1996 and December 2000, questionnaire booklets were included with the letters that are customarily sent to parents (or caregivers) of children who have been referred to the outpatient clinic to invite them for their initial visits. The booklets contained the CSBQ and the CBCL. Parents were asked to complete the questionnaires and bring them to the initial visit. They were informed that a psychiatrist could help with any questions.

At the clinic, child and adolescent psychiatrists carried out DSM-IV classification following extensive diagnostic procedures. These procedures included several clinical interviews, in which parents or caregivers were asked to describe the developmental history of their children and their present functioning in a variety of developmental domains. Play sessions with each child provided additional information about the child's social functioning, communication abilities, and imaginative abilities. School officials were asked to provide information about the children's behavior at school. Additional psychological assessment was conducted on approximately ten percent of the children to provide information about their cognitive abilities and/or other specific areas of functioning (e.g., attention and memory). Luteijn and colleagues (2000a) described data from a subsample of this group ($N = 916$), and this subsample was used to derive the original subscales of the CSBQ.

It is important to note that this diagnostic procedure led to clinical diagnoses; clinicians weighted the diverse sources of information according to expert judgment, and not on the basis of standardized algorithms. Parts of the procedure were therefore likely to differ from clinician to clinician. In contrast to standardized diagnostic interviews, which are designed to reduce such variability, this procedure introduced an unknown amount of error into the clinical classification of the children. While this situation imposes limitations on the present study, at least two factors made it infeasible to confirm the clinical diagnoses using standardized diagnostic interview. First, the number of children included in this study would have made such diagnostic verification daunting, if not impossible. For example, administration of both the Diagnostic Interview Schedule (Shaffer *et al*., 2000) and the ADI-R (Lord *et al*., 1994) in order to cover a broad range of disorders, including autism, can take up to five hours. Secondly, and more importantly, no standardized diagnostic interviews exist for the diagnostic group that was our main focus (i.e., PDDNOS). Our efforts to address problems related to diagnostic variation were therefore limited to inclusion of only those children who could be unambiguously classified while carefully excluding children with uncertain or dual diagnoses from all comparisons across clinical groups. The diagnostic procedure described above resulted in the following diagnostic groups:

High Functioning Autism (n = 102)

Only 5% of the children diagnosed with High Functioning Autism (HFA group) were recruited through the outpatient clinic. Most of the children (95%) in this group had been contacted through "Autism Teams," which specialize in the assessment and treatment of these children. Because of the scarcity of high functioning autistic children, the children in this group came from different parts of the Netherlands. The Autism Teams selected children who met the criteria for Autistic Disorder and whose IQ was higher than 70.

PDDNOS (n = 544)

The PDDNOS group consisted of children whose problems with social interaction, communication, or both were severe enough to have a negative impact on daily functioning. Many of them had restricted repertoires of activities and interests. None met the DSM-IV criteria for Autistic Disorder, Asperger's Disorder or other specific PDD categories. The problems of these children were therefore classified as PDDNOS. Comorbidity with other DSM-IV disorders occurred in 7.5% of the children and included the following problems: elimination disorders (e.g., enuresis, encopresis) in 4.2%, tic disorders in 1.5%, anxiety disorders in 1.1%, mood disorders in 0.4%, and eating disorders in 0.4% of the children. Note that comorbidity with ADHD and mental retardation are not included in these estimates and are treated as separate groups below.

ADHD (n = 585)

The ADHD group consisted of children with attention deficits, with or without hyperactivity problems. Oppositional defiant problems were allowed to be present in this group to the extent that they did not reach the clinical threshold for a diagnosis of Oppositional Defiant Disorder. Other forms of comorbidity occurred in 8.9% of the children in this group. These included: elimination disorders (e.g., enuresis, encopresis) in 5.3%, tic disorders in 1.5%, anxiety disorders in 1.4%, and mood disorders in 0.7% of the children.

Combined ADHD and PDDNOS (n = 164)

The children in this combined group met the DSM-IV criteria for both ADHD and PDDNOS. This group is referred to as the ADHD + PDDNOS group. Oppositional defiant problems were allowed to be present in this group to the extent that they did not reach the clinical threshold for a diagnosis of Oppositional Defiant Disorder. None of these children had significant additional psychiatric problems.

Internalizing Disorders (n = 209)

The children or adolescents in this group had anxiety disorders (41.6%), mood disorders (25.8%), somatization disorders (19.1%), and elimination disorders (13.4%). Comorbidity occurred in 3.8% of these children and concerned either another internalizing disorder (e.g., elimination disorders; 1.9%) or other problems (e.g., selective mutism or parent–child relational problems; 1.9%).

Other Psychiatric Problems (n = 666)

The remaining children in Subsample 1 could not be classified in any of the five categories described above. One of the following situations applied to these children: (1) they had been classified

under different diagnostic categories (e.g., learning disorders, tic disorders, sleep disorders, ODD, and CD), but the limited number of children in each of these diagnostic categories did not justify separate analyses; (2) they had not received Axis 1 diagnoses because of subthreshold psychopathology, or because they had been referred primarily as a result of parent–child relational problems or child abuse; (3) they had received dual diagnoses in which neither category of problems was clearly dominant. Data from this group of 666 children were used in the factor analytic studies but not in the clinical group comparisons.

Subsample 2 (N = 904) included Mentally Retarded Children, with (n = 152) or without (n = 581) PDD, a "Doubtful" PDD Group (n = 94), a Group with No Information on PDD (n = 11), and Children whose Intellectual Functioning was Described as "Borderline" (n = 66)

Mental retardation frequently occurs along with PDD (Bryson, 1997; Steffenburg & Gillberg, 1986; Wing, 1981; Wing & Gould, 1979). Should the CSBQ prove a suitable instrument for assessing PDDNOS symptoms across the whole range of intellectual functioning, this could add to its value.

As part of a total population-based investigation of PDD in mentally retarded children (between the ages of 4 and 18) in Friesland (a province in the northern part of the Netherlands), De Bildt and colleagues (2003) recruited CSBQ data from mentally retarded (IQ score at or below 70; n = 838) children and adolescents and from those functioning at a borderline intellectual level (IQ score above 70 and below 85; n = 66).

The children's parents were contacted through various facilities (e.g., schools for children with [severe] learning problems, day-care centers, and other institutions). Parents who were willing to participate in the study were contacted by mail or telephone, and they received questionnaire booklets that included the CSBQ and the CBCL. The questionnaires were followed by extensive interviewing, which included the Vineland Adaptive Behavior Scales (VABS, Sparrow, Balla, & Cicchetti, 1984).

Diagnoses of mental retardation or borderline intellectual functioning was based on standardized IQ tests, available from the care facilities. Information for 20% of the children was incomplete or absent, in which case diagnoses of mental retardation were based on clinical reviews of mental functioning, combined with assessments of the level of adaptive behavior as measured by the VABS.

The "Scale for Pervasive Developmental Disorder in individuals with Mental Retardation" (PDD-MRS) (Kraijer, 1997) was used to determine whether a child had a PDD. This is a Dutch instrument that has been widely studied and applied in the care of mentally retarded individuals in the Netherlands and Belgium (Kraijer, 1997; Kraijer & De Bildt, 2005). Sensitivity and specificity compared to clinician-assigned DSM-III-R diagnoses of PDD were 92.7 and 94.4, respectively (Kraijer, 1997). Scores of the PDD-MRS distinguish three categories (i.e., non-PDD, doubtful-PDD, and PDD). School psychologists or facility staff members completed PDD-MRS instruments for the children in this study. Using the instrument's classification system, researchers divided the mental retardation group into a PDD group (n = 152; referred to as the MR + PDD group), a non-PDD group (n = 581; referred to as the MR group) and a doubtful-PDD group (n = 94). For 11 children, the PDD-MRS data were missing. Because the PDD-MRS does not differentiate between the various PDDs, the MR + PDD group in the present study consists of children with autistic disorder as well as children with PDDNOS. Data from the MR + PDD and the MR groups were used in the clinical group comparisons. All 904 children in Subsample 2 were included in the factor analytic studies.

Subsample 3 (n = 232) Included typically Developing Children

Parents or caregivers of these children were approached through randomly selected elementary schools in the north of the Netherlands. Parents who were willing to participate received booklets containing the CSBQ and CBCL. The parents or caregivers of these children all declared that their children had never been in contact with psychological or psychiatric services and had not suffered from severe behavioral or emotional problems. This group of typically developing children is referred to as the normal control (NC) group.

General Characteristics of the Subsamples

The general characteristics of the subsamples included in the comparison of the clinical groups are summarized in Table I. The groups differed significantly with respect to mean age ($F(7, 2265) = 65.50$; $p < .001$). Children in the MR group were the oldest, while the children in the PDDNOS + ADHD group

Table I. General Characteristics of the Subsamples

	PDDNOS $N=544$	ADHD $N=586$	PDDNOS + ADHD $N=164$	HFA $N=102$	MR + PDD $N=152$	MR $N=581$	ID $N=209$	NC $N=232$
Age								
Range	4–18	4–18	4–16	4–18	3–18	4–18	4–18	4–14
Mean	8.99	8.84	7.85	9.42	11.22	11.98	10.90	8.34
SD	3.47	3.35	2.73	4.08	3.88	3.59	3.46	2.67
Gender								
Male (%)	78	86	87	85	76	59	49	45
Female (%)	22	14	13	15	24	41	51	55

were the youngest. The male–female ratio differed across diagnostic groups (Chi-square = 278.63, df = 7, $p < .001$). The PDDNOS, ADHD, PDD-NOS + ADHD, HFA, MR and MR + PDD groups contained more boys than girls (all six binomial tests: $p < .001$). This is consistent with the male–female ratios in prevalence and referral rates as reported in the literature for these diagnostic groups (Rutter, Caspi, & Moffit, 2003). The ID group and the NC group had equal male–female ratios.

Instruments

Children's Social Behavior Questionnaire (CSBQ)

In its original form the CSBQ (Luteijn *et al.*, 1998, 2000a) is a 96-item questionnaire for parents or caregivers of children with PDD. The items describe a broad range of severe and less severe behavioral features that are typical of PDD. They were formulated on the basis of the literature, parental descriptions, and clinical insights developed at the outpatient clinic for child and adolescent psychiatry in Groningen, the Netherlands. Parents are asked to indicate, on the basis of the child's behavior during the preceding two months, whether the behavior "does not apply" to the child (score 0), "sometimes or somewhat applies" (score 1), or "clearly or often applies" (score 2). The items are divided among five subscales, referring to "acting-out problems," "social contact problems," "social insight problems," "anxious/rigid" behaviors and "stereotypical" behaviors. Internal consistency reliability was $>.75$ for all subscales; Inter-rater reliability was $>.70$ for all subscales, except for the anxious/rigid subscale (ICC = .64); Test–retest reliability was $>.80$ for all subscales, except for the subscale "social insight problems" (ICC = .62) and the subscale "stereotypical" (ICC = .32) (Luteijn *et al.*, 2000a). Evidence for convergent validity of the subscales came from relatively high correlations with substantively similar subscales of the Child Behavior Checklist (Achenbach, 1991) and the Autism Behavior Checklist (Krug, Arick, & Almond, 1980).

The Child Behavior Checklist (CBCL)

The CBCL (Achenbach, 1991) was included in this study for the purpose of investigating the stability of the CBSQ factor structure and the convergent and divergent validity of the CSBQ subscales (described further below). The CBCL consists of 120 items that refer to diverse problem behaviors and emotions, as reported by parents based on the preceding six months. Research with the Dutch version of the CBCL has shown that results obtained in the Netherlands are comparable to those that have been obtained in the United States (Verhulst *et al.*, 1996). The items are rated using a 3-point scale, with 0 indicating "not true," 1 indicating "somewhat or sometimes true," and 2 indicating " very true or often true." Eighty-five of these 120 items are used in eight syndrome scales: "Withdrawn," "Somatic complaints," "Anxious/depressed," "Delinquent behavior," "Aggressive behavior," "Thought problems," "Social problems," and "Attention problems" (Achenbach, 1991). These 85 items were factor analyzed simultaneously with the CSBQ items. While the CSBQ focuses on the problem domain of PDD with dense item sampling within this domain, the CBCL measures across the broad range of childhood psychiatric problems. These characteristics make the CBCL an appropriate instrument for investigating both the convergent and divergent validity of the CSBQ.

Analyses

Selection of Items by Clinical Judgment

The first step in refining the CSBQ was to select only those items from the 96-item version of the CSBQ that were considered most characteristic of PDD. This selection on the conceptual basis preceded

the statistical analyses, as the outcome of a factor analysis (i.e., the basis for constructing the CSBQ subscales) may be highly dependent upon the items included in the analysis. The original CSBQ items had been chosen to be as broad and comprehensive as possible, in order to capture all possible behaviors observed in children with PDD. This was done despite the fact that some of these behaviors belong predominantly to different diagnostic categories. This broad focus of the item selection resulted in subscales that were also broad (i.e., heterogeneous) in focus, at the cost of precision in capturing the behaviors that were most relevant for the PDDs. Refining the CBSQ such that it consists of subscales that are specifically targeted towards the measurement of PDD but not other domains of psychopathology required screening the items for their relevance to PDD. On the basis of clinical judgment by the authors of this paper, 68 of the 96 items were selected for further statistical analysis. Inter-rater disagreements concerning which items to retain were minimal. Following discussion, the final set of items to be retained was determined by consensus. Items were removed if one of the followings applied: (a) they were very general and not specifically tied to psychopathology (e.g., ''pronounces words unclearly''); (b) they described typical ADHD symptoms (e.g., ''cannot sit still, some part of him/her is always moving''); or (c) they described worries (e.g., ''is over-concerned that something might happen to father/mother'').

Factor Analyses

A second selection of items was made on the basis of factor analysis. The size and diversity of the sample of children, who exhibited the full range of problems from normal to autistic, as well as from various diagnostic categories, ensured sufficient variance for the items. This enhanced the likelihood that findings would be replicable in future samples. Low item variance due to skewed ordinal score distributions, typical in the measurement of psychopathology, results in unstable correlation coefficients, which in turn leads to unstable factor structures, particularly when sample sizes are modest (Hartman *et al.*, 1999, 2001). For this reason, the complete sample was used in this study ($N = 3{,}407$).

Factor analysis was conducted using the maximum likelihood (ML) estimation method with Promax rotation (SPSS version 11.0). ML is the most commonly used method of estimation and has therefore been subject to the most thorough investigation. Promax rotation was selected, as it allows for correlations between the factors. The subscales of the CSBQ are likely to correlate with one another, as the various problem domains within PDD rarely occur in isolation. The factor analyses were conducted on a Pearson correlation matrix.

The number of factors was determined according to the scree-test (Cattell, 1966) and the eigenvalue-greater-than-one rule (Kaiser, 1960), as well as the substantive meaning of the factors. The factor loadings were subsequently evaluated on the basis of two criteria. First, we determined how characteristic an item was for its subscale, using a minimum loading of .350 as a criterion for inclusion in a subscale. In addition, we required items to show adequate specificity for their subscales, using a difference of greater than .150 between the main loading and possible secondary loading on other factors as a criterion for inclusion in a subscale. Items with low factor loadings ($< .350$) or items that loaded on two or more factors (differences between factor loadings $\leq .150$) were thus removed from the item pool. The selected items were subsequently factor analyzed to determine whether the factor structure persisted in this subset of items, and whether the items still fulfilled the item-selection criteria. Additionally, we examined the robustness of the factor structure for different age groups and levels of functioning.

The outcome of a factor analysis may depend heavily on the items that are included in the analysis. We therefore used a simultaneous factor analysis of the pooled items of the CSBQ and the CBCL to investigate the stability of the CSBQ factor structure (Achenbach, 1991). This analysis also shows how the dimensions of the CSBQ, if replicated, merge with similar domains (convergent validity) and are separate from different problem domains (divergent validity). This provides insight into the added value of the CSBQ for describing PDD-related problem behavior, as compared to such broadband instruments as the CBCL. The fact that both questionnaires use the same rating-scale format (i.e., short descriptions of the intended behavior followed by the response categories 0, 1, and 2) facilitated the simultaneous analysis of CSBQ and CBCL items. The same methods and criteria were applied in this analysis. The number of factors was determined on both of the statistical criteria mentioned above, and according to substantive criteria. The analyses were based on data from the 3,234 children for whom both the CSBQ and the CBCL had been completed.

Intercorrelations of CSBQ Subscales

In order to study the patterns of association between the different PDD dimensions, we calculated correlations between subscales. The complete sample of 3,407 children was used for this analysis. Although each subscale of the CSBQ refers to a specific PDD problem domain, we expected to find substantial correlations between the subscales.

Reliability

Three forms of reliability (internal, inter-rater, and test–retest) were studied for the CSBQ total scale and for the six subscales. Information from the complete sample of 3,407 children was used to determine internal consistency reliability, which was measured using Cronbach's α. Both test–retest reliability and inter-rater reliability were based on a subsample of children who had been consecutively referred to the outpatient clinic for child and adolescent psychiatry in Groningen, the Netherlands, during two periods in time: 1996 (approximately one-third of the sample) and 2001 (approximately two-thirds of the sample). The average age of the children in the complete sample was 8.87, with a standard deviation of 2.77. To determine inter-rater reliability, both mothers ($n = 70$) and fathers ($n = 70$) completed the CSBQ simultaneously and independently. Inter-rater reliability was established by means of the Intraclass Correlation Coefficient (ICC). To examine test–retest reliability, 59 mothers completed the CSBQ for the second time after an interval of approximately four weeks. The average age of the children in this subsample was 9.03, with a standard deviation of 2.87. Test–retest reliability was established by means of Pearson *r*.

Comparison of CSBQ Scores in the different Clinical and Control Groups

We studied the following (clinical) groups: HFA, PDDNOS, ADHD, ADHD + PDDNOS, ID, MR, MR + PDD, and NC. As described above, only those children that could be relatively unambiguously classified were included in these group comparisons.

If it is to be valuable as a clinical and research tool, the CSBQ should provide different score profiles for PDDNOS and variants of typical development (i.e., PDDNOS vs. NC), as well as for the more severe manifestations of PDD from the milder forms of PDD (i.e., HFA vs. PDDNOS). In addition, the applicability of the instrument would be enhanced by the ability to show different score profiles for mentally retarded children who do and do not exhibit PDD. Further, the instrument should provide different score profiles for PDDNOS and other child psychiatric problems, including ID and ADHD. A final issue concerns comparisons of the score profile of the ADHD + PDDNOS group with those of the relatively pure PDDNOS and ADHD groups. The question of whether a combined diagnosis of ADHD and PDD results in higher scores on the CBSQ and its subscales as compared with ADHD and PDD-NOS, or whether the scores are similar to either PDD or ADHD is of particular interest.

Analyses of variance (ANOVA) with *post hoc* comparisons between the groups were conducted on the total scale and the subscales. Differences were considered statistically significant if $p < .01$. In addition to these tests for statistical significance, we calculated effect sizes (Cohen's *d*), in order to evaluate the substantiality of the differences (Cohen, 1988, p. 10). This procedure was especially appropriate for the current study, as the sample size—and therefore the power to find statistically significant differences—differed considerably across clinical groups. The magnitude of an effect size does not depend on sample size and allows for more unambiguous interpretation. An effect size of .20 is characterized as small, an effect size of .50 as medium, and an effect size of .80 and above as large (Cohen, 1988, pp. 24–26).

RESULTS

Factor Analyses

Construction of New Subscales

We conducted a factor analysis of the 68 CSBQ items. The scree-plot began to level off after the first six eigenvalues. The eigenvalue-greater-than-one rule indicated that 13 factors should be retained. The latter rule is known to overestimate the number of factors under the following conditions: (1) 40 or more variables are included in the analysis, and (2) communalities are low, as is usually the case for item-level analyses (Stevens, 2002). Both conditions apply to the present situation. In addition to using these statistical rules, we established the number of factors on the basis of the substantive content of the factors, determining whether increasing the number of factors still allowed the items of a factor to measure a clinical concept, beginning with two factors and increasing to as many as 13. The latter consideration allowed six

conceptually different and meaningful factors. When seven or more factors were extracted, one or two items split off from the main factor to form either non-interpretable subscale(s) (i.e., none of the items had a factor loading ≥.35) or subscales based on very narrow content. Since further refinement was not substantively defensible, we decided to retain the six factors (subscales) indicated by the scree plot. Together, these factors accounted for 43.64% of the variance. The progressive differentiation into more homogeneous constructs from two to six factors is illustrated in Figure 1.

Having decided on the number of factors, we applied the two item-selection criteria described earlier (i.e. main loading ≥.35; difference between factor loadings >.15). The difference-between-loadings criterion resulted in the elimination of some items that described typical PDD symptoms but whose contributions were not specific to any of the six subscales. For example, the item "does not understand that certain things are ‚not done"' loaded high on both the "behavior/emotions not optimally tuned to the social situation" and the "orientation problems in time, place, or activity" subscales. It was therefore not included in the revised version of the CSBQ. Following this procedure, 49 items were retained and once again analyzed for factor structure. The factor structure for the reduced list was the same as that of the 68-items, and the item-selection criteria still applied for all 49 items. The six-factor solution accounted for 48.91% of the variance. An overview of the six subscales, the selected items, and the factor loadings is provided in Table II.

We examined the robustness of the factor structure for different age groups and levels of functioning. The factor solution was robust for both older (above the median age of nine years) and younger (nine years or younger) children, with each item having its highest loading on the appropriate factor. Similarly, the factor structure was robust when considering the children with and without MR separately. In these four more homogeneous subsamples, a few items did not meet the item-selection criteria due to reduced item variance.

CSBQ factor Structure among CBCL Items

We conducted a factor analysis on the 49 CSBQ and 85 CBCL items. The scree plot began to level off after the first nine eigenvalues. The eigenvalue-greater-than-one rule indicated the retention of 15 factors. When 10 factors were retained, the tenth factor was comprised of two items, which were very similar in content and had been part of the first factor in the nine-factor solution. The retention of this tenth factor added no substantive value to the model, as the two items fit in perfectly with the remaining items of factor one. In contrast, the nine-factor solution was clearly interpretable and resulted in nine conceptually different

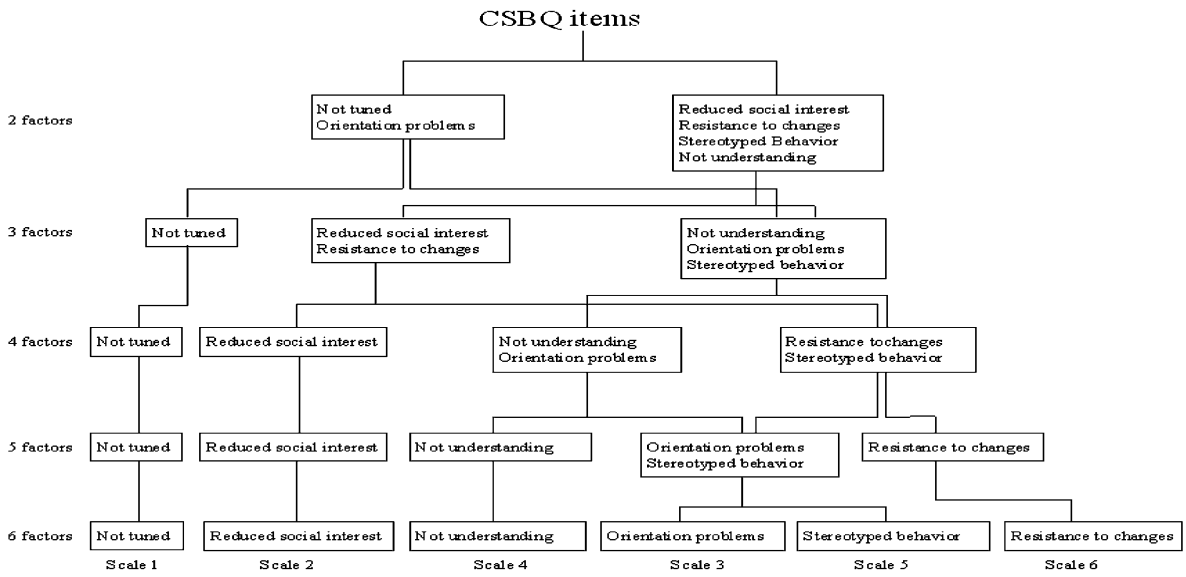

Fig. 1. Schematic representation of the outcome of the factor analyses.

Table II. Factor Structure of the CSBQ

	Factor					
	1	2	3	4	5	6
31 Quickly gets angry	.832					
43 Does not know when to stop, e.g., goes on and on about things	.713					
44 Is extremely stubborn	.670					
32 Stays angry for a long time e.g., when he/she does not get his/her way	.663					
37 Is disobedient	.656		.266	−.184		
9 Draws excessive attention to him/herself	.637	−.180				
30 Shows sudden changes of mood	.630					
42 Makes a fuss over little things; "makes a mountain of a mole-hill"	.627		−.233	.201		
8 Over-reacts to everything and everyone	.604	−.203				
38 Cannot be corrected in situations in which he/she has done something wrong	.578	.157	.260			
40 Makes inconsiderate remarks e.g., remarks that are painful to others	.462			.267		
20 Has little or no need for contact with others		.646	−.183			
16 Makes little eye contact		.615	.249			
18 Does not seek comfort when he/she is hurt or upset		.608				
17 Dislikes physical contact e.g., does not want to be touched or hugged		.567				
21 Does not respond to initiatives by others e.g., does not play along when asked		.564			.181	
19 Does not initiate play with other children		.554	−.174			
14 Acts as if others are not there		.549				
15 Lives in a world of his/her own		.525	.157			
34 Does not show his/her feelings in facial expressions and/or bodily posture		.512				
13 Does not look up when spoken to		.500	.322			
33 Cannot be made enthusiastic about anything; does not particularly like anything	.169	.425				
41 Does not appreciate it when someone else is hurt or sad	.211	.418				
29 Does things without realizing the aim e.g., constantly has to be reminded to finish things			.664			
28 Does things without realizing what stage of the activity he/she has reached (beginning, middle, ending)			.647			
49 Has no sense of time			.630			
39 Takes in information with difficulty			.605			
27 Has difficulties doing two things simultaneously e.g., he/she cannot dress and listen to parent at the same time			.596			
35 Does not appreciate danger			.541			
48 Gets lost easily e.g., when out with someone			.503			
36 Barely distinguishes between strangers and familiar people e.g., readily goes with strangers			.387		.217	
6 Takes things literally e.g., does not understand certain expressions				.821		
5 Does not understand jokes				.782		
3 Does not fully understand what is being said to him/her i.e., tends to miss the point			.204	.655		
7 Is exceptionally naive; believes anything you say				.654		
4 Frequently says things that are not relevant to the conversation				.574		
1 Talks confusedly; jumps from one subject to another in speaking			.201	.425		
2 Only talks about things that are of concern for himself/herself	.211			.364		
25 Constantly feels objects					.566	
24 Smells objects					.536	
11 Makes odd, fast movements with fingers or hands					.520	
23 Is extremely pleased by certain movements and keeps doing them e.g., turning around and around					.515	
10 Flaps arms/hands when excited					.509	
26 Is fascinated by certain colors, forms, or moving objects					.508	
12 Sways to and fro					.433	
22 Is unusually sensitive to certain sounds (e.g., always hears certain sounds earlier than other people					.368	.162
46 Remains clammed up in new situations or if change occurs						.839
45 Panics in new situations or if change occurs						.832
47 Opposes change						.720

Note: Items are arranged by descending factor loading; loadings ≤ .150 are not reported; n = 3407; Factor 1: behavior/emotions not optimally tuned to the social situation; Factor 2: reduced contact and social interest; Factor 3: orientation problems in time, place, or activity; Factor 4: difficulties in understanding social information; Factor 5: stereotyped behavior; Factor 6: fear of and resistance to changes.

constructs. We therefore retained nine factors, which together accounted for 41.59% of the variance. Table III provides an overview of these factors. The table shows only the items that fulfilled both item-selection criteria.

Two factors, ''difficulties in understanding social information'' and ''fear of and resistance to changes'', consisted solely of items from the CSBQ. The CSBQ ''stereotyped behavior'' factor included one additional item from the CBCL, but remained conceptually identical. Three factors consisted of uniquely CBCL items: ''delinquency'', ''social problems (in the sense of being teased or disliked by others)'', and ''internalizing problems''.

Three factors combined CSBQ and CBCL items. First, the CSBQ subscale ''behavior/emotions not optimally tuned to the social situation'' was complemented with items from the CBCL ''aggression'' subscale. Second, the CSBQ subscale ''orientation problems in time, place, or activity'' combined with the CBCL items ''can't concentrate'', ''poor schoolwork'', and ''clumsy''. Third, the CSBQ subscale ''reduced contact and social interest'' combined with five items from the CBCL ''withdrawn'' subscale.

In summary, the six factors of the CSBQ reemerged in the simultaneous factor analysis of CSBQ and CBCL items. This argues for the stability of the CSBQ factor structure when analyzing a much broader

Table III. Nine-factor Solution from Simultaneous Factor Analysis of CSBQ and CBCL Items

Factor 1: Behavior/emotions not optimally tuned to the social situation /aggressive behavior (10 CSBQ and 17 CBCL items)		
CSBQ 08. Over-reacts to everything and everyone	CBCL 03. Argues	CBCL 86. Stubborn
CSBQ 09. Draws excessive attention to him/herself	CBCL 10. Can't sit still	CBCL 87. Mood changes
CSBQ 30. Shows sudden changes of mood	CBCL 16. Mean to others	CBCL 88. Sulks
CSBQ 31. Quickly gets angry	CBCL 19. Demands attention	CBCL 90. Swearing
CSBQ 32. Stays angry for a long time	CBCL 22. Disobedient home	CBCL 93. Talks too much
CSBQ 37. Is disobedient	CBCL 26. No Guilt	CBCL 94. Teases
CSBQ 38. Cannot be corrected	CBCL 27. Jealous	CBCL 95. Temper tantrums
CSBQ 42. Makes a fuss over little things	CBCL 68. Screams	CBCL 104. Loud
CSBQ 43. Doesn't know when to stop	CBCL 74. Shows off	
CSBQ 44. Is extremely stubborn		
Factor 2: Internalizing problems (20 CBCL items)		
	CBCL 09. Can't get mind off things	CBCL 52. Guilty
	CBCL 12. Lonely	CBCL 54. Overtired
	CBCL 31. Fear do bad	CBCL 56a. Aches
	CBCL 32. Be perfect	CBCL 56b. Headaches
	CBCL 33. Unloved	CBCL 56c. Nausea
	CBCL 35. Worthless	CBCL 56f. Stomach
	CBCL 40. Hears things	CBCL 71. Self-conscious
	CBCL 45. Nervous	CBCL 85. Strange ideas
	CBCL 50. Fearful	CBCL 103. Sad
	CBCL 51. Dizzy	CBCL 112. Worries
Factor 3: Reduced contact and social interest/withdrawn (11 CSBQ and 6 CBCL items)		
CSBQ 13. Does not look up when spoken to	CBCL 42. Rather be alone	
CSBQ 14. Acts as if others are not there	CBCL 65. Refuses talk	
CSBQ 15. Lives in a world of his/her own	CBCL 69. Secretive	
CSBQ 16. Makes little eye contact	CBCL 80. Stares blankly	
CSBQ 17. Dislikes physical contact	CBCL 102. Underactive	
CSBQ 18. Does not seek comfort	CBCL 111. Withdrawn	
CSBQ 19. Does not initiate play with other children		
CSBQ 20. Has little/no need for contact		
CSBQ 21. Doesn't respond to initiatives by others		
CSBQ 33. Cannot be made enthusiastic		
CSBQ 34. Does not show his/her feelings face/body		
Factor 4: Delinquent behavior (7 CBCL items)		
	CBCL 21. Destroys other	
	CBCL 39. Bad companions	
	CBCL 72. Sets fire	
	CBCL 81. Steals at home	
	CBCL 82. Steals outside	
	CBCL 97. Threatens	
	CBCL 106. Vandalism	

Table III. (Continued)

Factor 5: Difficulties in understanding social information (7 CSBQ items)	
CSBQ 01. Talks confusedly	
CSBQ 02. Only talks about things that are of concern for himself/herself	
CSBQ 03. Does not fully understand what is said	
CSBQ 04. Says things not relevant to the conversation	
CSBQ 05. Does not understand jokes	
CSBQ 06. Takes things literally	
CSBQ 07. Is exceptionally naive	
Factor 6: Stereotyped behavior (8 CSBQ and 1 CBCL item)	
CSBQ 10. Flaps arms/hands when excited	CBCL 66. Repeats acts
CSBQ 11. Odd, fast movements with fingers or hands	
CSBQ 12. Sways to and fro	
CSBQ 22. Is unusually sensitive to certain sounds	
CSBQ 23. Pleased by certain movement	
CSBQ 24. Smells objects	
CSBQ 25. Constantly feels objects	
CSBQ 26. Is fascinated by colors, forms, moving objects	
Factor 7: Orientation problems in time, place, or activity (5 CSBQ and 3 CBCL items)	
	CBCL 08. Can't concentrate
CSBQ 28. Does not realize what stage of the activity	CBCL 61. Poor school work
CSBQ 29. Does things without realizing the aim	CBCL 62. Clumsy
CSBQ 39. Takes in information with difficulty	
CSBQ 49. Gets lost easily	
Factor 8: Fear of and resistance to changes (3 CSBQ items)	
CSBQ 46. Remains clammed up	
CSBQ 45. Panics	
CSBQ 47. Opposes change	
Factor 9: Social problems (3 CBCL items)	
	CBCL 25. Does not get along
	CBCL 38. Teased
	CBCL 48. Not liked

Note: The list includes only those items that had a main factor loading ≥.35, and a difference > .15 with possible secondary factor loadings; $n = 3234$. Both CSBQ and CBCL items are abbreviated.

item pool. Additionally, to the extent that the CSBQ subscales were complemented with items from the CBCL, the content of the six CSBQ subscales remained very similar, which is supportive of convergent validity. Divergent validity was also as expected. That is, consistent with its intended scope of measurement, the CSBQ does not measure "delinquency", "social problems (in the sense of being teased or disliked by others)", or "internalizing problems".

Intercorrelations of CSBQ Subscales

Table IV shows the correlation matrix for the six subscales of the CSBQ. The correlations between the subscales varied from .32 to .59.

Reliability of the Revised Subscales

Internal Consistency Reliability

Based on the minimum standard for reliability of .70 (Nunnally & Bernstein, 1994, p. 265), the internal consistency (Cronbach's α) of the total scale was good (.94 for 49 items). The internal consistencies of the subscales were also good: Subscale 1 "not optimally tuned to the social situation" ($\alpha = .90$); Subscale 2 "reduced contact and social interest" ($\alpha = .85$); Subscale 3 "orientation problems in time, place, or activity" ($\alpha = .84$); Subscale 4 "difficulties in understanding social information" ($\alpha = .85$); Subscale 5 "stereotyped behavior" ($\alpha = .76$); and Subscale 6 "fear of and resistance to changes" ($\alpha = .85$).

Inter-rater Reliability

Inter-rater reliability was good for the CSBQ total scale (ICC = .86) as well as for the subscales: "not optimally tuned to the social situation" (ICC = .89), "reduced contact and social interest" (ICC = .79), "orientation problems in time, place, or activity" (ICC = .81), "difficulties in understanding social information" (ICC = .87), "stereotyped

Table IV. Intercorrelations between CSBQ Subscales

Subscale	Tuned	Social	Orientation	Understanding	Stereotyped	Change
Tuned	–					
Social	.43	–				
Orientation	.52	.43	–			
Understanding	.49	.47	.59	–		
Stereotyped	.33	.40	.46	.39	–	
Change	.39	.41	.32	.37	.36	–

Note: $n = 3407$. All correlations differ significantly from zero ($p < .01$), tuned: behavior/emotions not optimally tuned to the social situation; social: reduced contact and social interest; orientation: orientation problems in time, place, or activity; understanding: difficulties in understanding social information; stereotype: stereotyped behavior; change: fear of and resistance to changes.

behavior'' (ICC = .75), and ''fear of and resistance to changes'' (ICC = .80).

Test–retest Reliability

For all scales, test–retest reliability was good: total score (r = .90); ''not optimally tuned to the social situation'' (r = .89); ''reduced contact and social interest'' (r = .88); ''orientation problems in time, place, or activity'' (r = .82); ''difficulties in understanding social information'' (r = .80); ''stereotyped behavior'' (r = .80), and ''fear of and resistance to changes'' (r = .83).

Comparison of CSBQ Scores in the different Clinical and Control Groups

Table V provides means and standard deviations for each diagnostic group and for each (sub)scale.

Pairwise comparisons of the different diagnostic groups provided the following results (see also *Note* in Table V). The PDDNOS group had significantly higher scores on the CSBQ total score and all of the CSBQ subscales, as compared to the NC group. Effect sizes were 2.13 for total score, 1.71 for ''behavior/emotions not optimally tuned to the social situation'', 1.70 for ''reduced contact and social interest'', 1.62 for ''orientation problems in time, place, or activity'', 1.08 for ''difficulties in understanding social information'', .89 for ''stereotyped behavior'', and 1.02 for ''fear of and resistance to changes''. All of these effects sizes are large, according to Cohen's criteria (Cohen, 1988, pp. 24–26).

Scores from the HFA group were significantly higher than those of the PDDNOS group on the CSBQ total scale (effect size .60) and on the following subscales: ''reduced contact and social interest'' (effect size .52), ''difficulties in understanding social information'' (effect size .77), ''stereotyped behavior'' (effect size .61), and ''fear of and resistance to changes'' (effect size .40). These effect sizes range from small to large (Cohen, 1988). No significant differences were found for the subscales ''behavior/emotions not optimally tuned to the social situation'' (effect size −.04) and ''orientation problems in time, place, or activity'' (effect size .34).

Total scores on the CSBQ were higher for the MR + PDD group than they were for the MR group (effect size .76), and the groups differed significantly on the following subscales: ''reduced contact and social interest'' (effect size .84), ''orientation problems in time, place, or activity'' (effect size .62), ''stereotyped behavior'' (effect size .97), and ''fear of and resistance to changes'' (effect size .39). These effect sizes range from small to large (Cohen, 1988). The groups did not differ on the ''behavior/emotions not optimally tuned to the social situation'' (effect size .25) and ''difficulties in understanding social information'' (effect size .05) subscales.

In comparison with the ID group, children in the PDDNOS group had significantly higher scores on the CSBQ as a whole and all of the CSBQ subscales. Effect sizes were 1.35 for total score, .91 for ''behavior/emotions not optimally tuned to the social situation'', .85 for ''reduced contact and social interest'', 1.33 for ''orientation problems in time, place, or activity'', 1.13 for ''difficulties in understanding social information'', .68 for ''stereotyped behavior'', and .40 for ''fear of and resistance to changes''. With the exception of the last one, all of these effect sizes are large (Cohen, 1988).

The PDDNOS group differed significantly from the ADHD group with regard to the CSBQ total score (effect size .37) and the following subscales: ''reduced contact and social interest'' (effect size .81), ''difficulties in understanding social information'' (effect size .33), ''stereotyped behavior'' (effect size .26), and ''fear of and resistance to changes'' (effect size .50). These effect sizes range from small to large

Table V. Means and Standard Deviations for each (sub)scale and Diagnostic Group

	Total			Tuned		Social		Orientation		Understanding		Stereotyped		Change	
	N	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
HFA	102	47.22	15.37	12.13	5.23	10.22	4.66	7.71	3.80	8.89	3.29	5.19	4.22	3.09	1.72
PPDDNOS	544	37.84	15.94	12.36	5.60	7.75	4.91	6.42	3.80	6.12	3.91	2.90	3.18	2.30	2.16
ADHD	586	32.22	14.08	12.63	5.15	4.17	3.90	7.05	3.62	4.89	3.57	2.15	2.59	1.33	1.66
ADHD + PDDNOS	164	42.66	14.51	14.99	4.50	5.76	4.29	9.09	4.13	6.98	3.60	3.51	3.17	2.34	2.12
ID	209	18.54	12.44	7.40	5.29	4.01	3.85	2.08	2.60	2.39	2.57	1.17	1.64	1.49	1.89
MR + PDD	152	33.64	15.71	7.03	4.79	7.53	5.42	7.41	4.20	4.71	3.84	5.09	4.06	1.87	1.86
MR	581	21.71	15.55	5.78	5.11	3.54	3.95	4.89	3.96	4.50	3.72	1.80	2.51	1.20	1.57
NC	232	10.28	9.05	4.24	3.71	1.38	1.99	1.45	2.10	1.94	2.13	.74	1.24	.55	1.13

Note: Means and standard deviations are based on raw scores.
Pairwise statistical tests: ,>': significantly higher score ($p < .01$); ,=': no significant difference ($p \geq .01$).
(1) Total = total CSBQ score: PDDNOS > NC; HFA > PDDNOS; MR + PDD > MR; PDDNOS > ID; PDDNOS > ADHD; PDDNOS + ADHD > PDDNOS; PDDNOS + ADHD > ADHD.
(2) Tuned = behavior/emotions not optimally tuned to the social situation: PDDNOS > NC; HFA = PDDNOS; MR + PDD = MR; PDDNOS > ID; PDDNOS = ADHD; PDDNOS + ADHD > PDDNOS; PDDNOS + ADHD > ADHD.
(3) Social = reduced contact and social interest; PDDNOS > NC; HFA > PDDNOS; MR + PDD > MR; PDDNOS > ID; PDDNOS > ADHD; PDDNOS > PDDNOS + ADHD; PDDNOS + ADHD > ADHD.
(4) Orientation = orientation problems in time, place, or activity; PDDNOS > NC; HFA = PDDNOS; MR + PDD > MR; PDDNOS > ID; PDDNOS = ADHD; PDDNOS + ADHD > PDDNOS; PDDNOS + ADHD > ADHD.
(5) Understanding = difficulties in understanding social information; PDDNOS > NC; HFA > PDDNOS; MR + PDD = MR; PDDNOS > ID; PDDNOS > ADHD; PDDNOS + ADHD = PDDNOS; PDDNOS + ADHD > ADHD.
(6) Stereotyped = stereotyped behavior; PDDNOS > NC; HFA > PDDNOS; MR + PDD > MR; PDDNOS > ID; PDDNOS > ADHD; PDDNOS + ADHD = PDDNOS; PDDNOS + ADHD > ADHD.
(7) Change = fear of and resistance to changes; PDDNOS > NC; HFA > PDDNOS; MR + PDD > MR; PDDNOS > ID; PDDNOS > ADHD; PDDNOS + ADHD = PDDNOS; PDDNOS + ADHD > ADHD.

(Cohen, 1988). The PDDNOS group differed most from the ADHD group in terms of "reduced social contact and interest". No significant differences were found on "behavior/emotions not optimally tuned to the social situation" (effect size −.05) and the "orientation problems in time, place, or activity" (effect size −.17) subscales.

Finally, we found that combined diagnoses of PDD and ADHD led to significantly higher scores on the total scale than did separate diagnoses of either PDDNOS or ADHD (effect sizes .32 and .73, respectively), as well as on the following subscales: "behavior/emotions not optimally tuned to the social situation" (effect sizes .52 and .49, respectively) and "orientation problems in time, place, or activity" (effect sizes .67 and .53, respectively). In contrast, scores of the combined PDD + ADHD group did not differ from scores of the PDDNOS group on the subscales "difficulties in understanding social information" (effect size .23), "stereotyped behavior" (effect size .19) and "fear of and resistance to changes" (effect size .02). On each of these three subscales, however, and also on the "reduced contact and social interest" scale, the combined ADHD + PDD group had significantly higher scores than did the ADHD group (effect sizes .58, .47, .53, and .39, respectively). The combined PDD + ADHD and PDD groups were therefore more alike than were the PDD + ADHD and ADHD groups. On the "reduced contact and social interest" scale, scores for the combined PDD + ADHD group were lower than those of the PDDNOS group (effect size −.43).

DISCUSSION

The goal of this study was to refine the subscales of the CSBQ and reduce the length of the instrument, while maintaining good psychometric properties. The findings of the study indicate that the revised instrument is both reliable and valid. The procedures described in this paper led to a more concise CSBQ, reducing the number of items from 96 to 49. Six homogeneous subscales were constructed, which allow for a differentiated description of PDD problems. The content of these revised subscales is similar to that of the subscales derived from the original 96-item version of the CSBQ (Luteijn *et al.*, 2000a). In

addition, this study has demonstrated the stability of the CSBQ factor structure by showing that the constructs of the CSBQ remained very similar under a simultaneous factor analysis of CSBQ and CBCL items. The six PDD problem domains that are differentiated by the CSBQ thus seem to be firmly anchored in the data.

Compared to the 96-item version, the current revised CSBQ gained in specificity for PDD, as problem items that tended to be more indicative of problem domains other than PDD (e.g., ADHD, worries) were removed. Specificity was further improved by removing items that were only marginally characteristic of a PDD dimension (i.e., those with low factor loading) and items that did not differentiate between different PDD dimensions (i.e., substantial loadings on more than one factor). In addition, the instrument was improved by allowing a more differentiated description of the former ''social insight problems'' subscale by dividing it into two more homogeneous subscales: ''orientation problems in time, place, or activity'' and ''difficulties in understanding social information''. All of these improvements enhance the ability to arrive at an unequivocal interpretation of the subscales of the CSBQ.

One important finding revealed by this study is that the reliability of the instrument did not suffer in any way from this reduction of items. The subscales showed good internal consistency and test–retest and inter-rater reliability estimates, all of which are either similar to or better than the reliability estimates for the 96-item version (Luteijn *et al*., 2000a). The outcome of simultaneous factor analyses of the SCBQ and the CBCL showed the expected patterns of convergent and divergent validity.

In further support of the validity of the CSBQ, the differences in subscale averages between the clinical and control groups were as expected. The scores of children with PDDNOS were higher than those of typically developing children for the CSBQ total scale, as well as for all of the subscales. The same result was found when comparing the PDD-NOS group to the ID group. Within the group of mentally retarded children, the CSBQ distinguished between the profiles of the PDD group and the non-PDD group, with the PDD group having higher scores. The total scores of children with PDDNOS were higher than were those of children with ADHD, as were their scores on all but two of the subscales. In contrast, the PDDNOS group had lower total CSBQ scores than did the HFA group, and its scores were also lower for all but two subscales. The latter finding is consistent with the idea that PDDNOS is a collection of disorders that share the core features of Autistic Disorder, although in a milder form (Buitelaar, Van der Gaag, Klin, & Volkmar, 1999; Towbin, 1997, Walker *et al*., 2004).

Scores on two subscales (''behavior/emotions not optimally tuned to the social situation'' and ''orientation problems in time, place, or activity'') were not consistent with the expectation that the problems of children with PDDNOS would be milder than those of children with HFA. These two subscales warrant further discussion. Not only did they not reveal differences between HFA and PDDNOS, they also failed to differentiate between PDDNOS and ADHD. These two subscales may be conceptually related to the ADHD dimensions of ''hyperactivity/impulsivity'' and ''attention problems'', respectively, as formulated in the DSM-IV-TR. An interesting finding in this respect is that a combined diagnoses of ADHD and PDD led to significantly higher scores than did separate diagnoses of either ADHD or PDD. This suggests that each condition (ADHD and PDD) has its own impact on the behaviors measured by these two subscales. For example, children with combined diagnoses of ADHD and PDD may, in addition to being very active and impulsive, have difficulty understanding what the appropriate behavior is in a given social context, which subsequently hinders their ability to adjust their behavior. It is therefore likely that the impairments that underlie these disorders, although unknown at this point, are separate mechanisms, each of which has its own impact. Together, these mechanisms add up or interact to produce more severe symptoms. While the underlying impairments may differ, the areas of overlap at the behavioral level between the PDDs and ADHD illustrate the difficulties that clinicians experience in classifying PDDNOS and ADHD (Jensen, Larrieu, & Mack, 1997; Barkley, 1997; Luteijn *et al*., 2000b).

In addition to the favorable results concerning the substantive and psychometric characteristics of the revised CSBQ, it is important to emphasize the practical gains resulting from reducing the number of items from 96 to 49. The streamlined instrument will be less costly and easier to administer, in both clinical and research settings.

The encouraging results reported above notwithstanding, three critical comments must be made. The present study was limited in that the clinical diagnoses were not confirmed with standardized diagnostic instruments. We must thus assume that, had such

assessments been applied, a number of children would have received diagnoses that differed from their current classifications. Given the number of children included in the sample, however, the task of conducting standardized diagnostic interviews would have been an enormous, if not impossible, task. More importantly, no standardized diagnostic interviews exist for the clinical group that was the central focus in this study (PDDNOS). As argued in the introduction, while an instrument such as the ADI-R is applicable for diagnosing Autism, it does not tap the subtler or less severe PDD problems required for a diagnosis of PDDNOS. In order to avoid faulty classifications, we opted for the strategy of excluding those children whose problems did not fit well into a single diagnostic category from the analysis. Nevertheless, the current classifications are certain to involve some degree of error, and the current results must therefore be seen as approximate to results based on standardized diagnostic instruments, had they been available.

A second critical comment concerns the differentiation of the subscales "orientation problems in time, place, or activity" and "difficulties in understanding social information." The substantive content of the "social insight problems" subscale of the 96-item questionnaire was one of the factors motivating the re-evaluation of the CSBQ. This subscale included items that referred to problems in understanding social information, in addition to items that referred to attention problems, a subset of which were part of the DSM-IV criteria for ADHD. This mixture of item content impeded the interpretation of high or low subscale scores. Although the revised 49-item version of the CBSQ (in which the ADHD symptoms have been removed) divides this dimension into two separate subscales—"orientation problems in time, place, or activity" and "difficulties in understanding social information"—their differentiation may be the least stable, as the correlation between them was the highest of all subscale inter-correlations ($r = .59$). The distinction between the two domains is nonetheless valuable, because, as discussed above, the former is part of the more specific PDD problems while the latter is part of the overlap in problems between PDD and ADHD. Future factor analytic study of the CSBQ in an independent sample will reveal the extent to which the current factor structure can be replicated and therefore how robust the distinction is.

A third critical comment concerns the subscale "fear of and resistance to changes," which consists of only three items. We decided to retain this subscale in the 49-item CSBQ, as resistance to changes is a fundamental characteristic of the PDDs. The content of the subscale is rather narrow, however, and should be extended in further efforts to develop the CSBQ.

We conclude by providing some examples of how the CSBQ can be used. For research purposes, the instrument facilitates the selection of samples that are more homogeneous, with respect to either the number of PDD problems (e.g., based on the severity of the CSBQ total score) or the type of problems (e.g., based on the relative severity profile produced by scores on the CSBQ subscales). Different dimensions within the PDD spectrum, as summarized by the CSBQ subscales, may relate differentially to underlying cognitive deficits, neurobiological parameters, or molecular genetic findings.

For clinical purposes, it is important to note that the CBSQ is not intended for purposes of diagnostic classification. It may nonetheless provide clinicians with a more specific and detailed description of the number and type of PDD(NOS) problems. The CSBQ might also be helpful as a first screening device when teachers or general practitioners suspect that a child has problems that fall within the spectrum of the PDDs. In these cases, the score profile of the child may help to plan for more specific diagnostic assessment and treatment, as well as their subsequent evaluation.

Finally, in both research and clinical settings, the CSBQ is useful for describing the severity and pattern of social deficits in clinical groups other than the PDDs. Our results showed that PDD-related problems can be present to varying degrees in a variety of child psychiatric disorders, and not just in those that fall under the umbrella of the autistic spectrum. The causal mechanisms underlying this finding are likely to be complex. Constantino and colleagues (2000) argued that most childhood psychiatric problems exert their negative influence on social behavior. Conversely, social deficits in PDD(NOS) may operate to exacerbate problems associated with other psychiatric conditions, as when a child with PDD, who is sensitive to change, responds in an oppositional manner to the disruption of daily routines (Constantino *et al.*, 2000). Further, an underlying causal mechanism may be operating, for example, when the same genes code for genetic variants that contribute to common phenotypic presentations across multiple clinical conditions, such as PDD and ADHD (Bakker *et al.*, 2003, Smalley, 1997). Charting the extent to which social problems are present, even when the problems suggest a diagnosis outside the PDD

spectrum, may often be useful and sometimes provide clues for a different primary diagnosis. The CSBQ is a psychometrically sound and easily administered tool for precisely this type of assessment.

REFERENCES

Achenbach, T. M. (1991). *Manual for the Child Behavior Checklist/4–18 and 1991 profile*. Burlington, VT: University of Vermont, Department of Psychiatry.

Althaus, M. (2000). Visual attention and autonomic adaptivity to attention-demanding tasks in children with autistic-type behavioural problems. PhD. thesis. University of Groningen.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Association.

Bakker, S. C., van der Meulen, E. M., Buitelaar, J. K., Sandkuijl, L. A., Pauls, D. L., Monsuur, A. J., van ,t Slot, R., Minderaa, R. B., Gunning, W. B., Pearson, P. L., & Sinke, R. J. (2003). A whole genome scan in 164 Dutch sib pairs with attention-deficit/hyperactivity disorder: Suggestive evidence for linkage on chromosomes 7p and 15q. *American Journal of Human genetics*, *72*, 1251–1260.

Barkley, R. A. (1997). *Attention deficit hyperactivity disorder: A handbook for diagnosis and treatment* (2nd ed.). New York: Guilford Press.

Beglinger, L. J., & Tristram, H. Smith. (2001). A review of subtyping in autism and proposed dimensional classification model. *Journal of Autism and Developmental Disorders*, *31*, 411–422.

Berument, S. K., Rutter, M., Lord, C., Pickles, A., & Bailey, A. (1999). Autism screening questionnaire: diagnostic validity. *British Journal of Psychiatry*, *175*, 444–451.

Bildt, A. de., Sytema, S., Kraijer, D., Ketelaars, C., Volkmar, F., & Minderaa, R. (2003). Measuring pervasive developmental disorders in children and adolescents with mental retardation. *Journal of Autism and Developmental Disorders*, *33*, 595–605.

Bishop, D. V., & Norbury, C. F. (2002). Exploring the borderlands of autistic disorder and specific language impairment: A study using standardized diagnostic instruments. *Journal of Child Psychology and Psychiatry*, *43*, 917–929.

Bryson, S. E. (1997). Epidemiology of autism: Overview and issues outstanding. In: D. J. Cohen, & F. R. Volkmar (Eds.), *Handbook of autism and developmental disorders*. (pp. 41–46). New York: John Wiley & Sons.

Buitelaar, J. K., Van der Gaag, R., Klin, A., & Volkmar, F. (1999). Exploring the boundaries of pervasive developmental disorder not otherwise specified: Analyses of data from the DSM-IV Autistic Disorder Field Trial. *Journal of Autism and Developmental Disorders*, *29*, 33–43.

Cattell, R. B. (1966). The meaning and strategic use of factor analysis. In: R. B. Cattell (Ed.), *Handbook of multivariate experimental psychology*. (pp. 174–243). Chicago: Rand McNally.

Chakrabarti, S., & Fombonne, E. (2001). Pervasive developmental disorders in preschool children. *Journal of the American Medical Association*, *285*, 3093–3099.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale: Lawrence Erlbaum.

Constantino, J. N., Przybeck, T., Friesen, D., & Todd, R. O. (2000). Reciprocal social behavior in children with and without pervasive developmental disorders. *Journal of Developmental and Behavioral Pediatrics*, *21*, 2–11.

Hartman, C. A. (2000). *Changing concepts of child psychopathology*. Dissertation, Amsterdam, The Netherlands: TT Publications.

Hartman, C. A., Hox, J., Auerbach, J., Erol, N., Fonseca, A. C., Mellenbergh, G. J., Nøvik, T. S., Oosterlaan, J., Roussos, A. C., Shalev, R. S., Zilber, N., & Sergeant, J. A. (1999). Syndrome dimensions of the Child Behavior Checklist and the Teacher Report Form: A critical empirical evaluation. *Journal of Child Psychology and Psychiatry*, *40*, 1095–1116.

Hartman, C. A., Hox, J., Mellenbergh, G. J., Boyle, M., Offord, D., Racine, Y., McNamee, J., Gadow, K., Sprafkin, J., Kelly, K. L., Nolan, E., Tannock, R., Schachar, R., Schut, H., Postma, I., Drost, R., & Sergeant, J. A. (2001). DSM-IV internal construct validity: When a taxonomy meets data. *Journal of Child Psychology and Psychiatry*, *42*, 817–836.

Jensen, V. K., Larrieu, J. A., & Mack, K. K. (1997). Differential diagnosis between attention-deficit/hyperactivity disorder and pervasive developmental disorder-not otherwise specified. *Clinical Pediatrics*, *36*, 555–561.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, *20*, 141–151.

Kraijer, D. W. (1997). *Autism and autistic-like conditions in mental retardation*. Lisse: Swets & Zeitlinger.

Kraijer, D. W., & de Bildt, A. (2005). The PDD-MRS: An instrument for identification of Autism Spectrum Disorders in persons with mental retardation. *Journal of Autism and Developmental Disorders*, *35*, 499–513.

Krug, D. A., Arick, J. R., & Almond, P. J. (1980). *Autism screening instrument for educational planning*. Portland, OR: ASIEP.

Lord, C., Rutter, M., & LeCouteur, A. (1994). Autism diagnostic interview-revised: A revised version of a diagnostic interview for caregivers of individuals with possible Pervasive Developmental Disorders. *Journal of Autism and Developmental Disorders*, *24*, 659–685.

Luteijn, E. F., Jackson, A. E., Volkmar, F. R., & Minderaa, R. B. (1998). The development of the Children's Social Behavior Questionnaire: Preliminary data. *Journal of Autism and Developmental Disorders*, *28*, 559–565.

Luteijn, E. F., Luteijn, F., Jackson, A. E., Volkmar, F. R., & Minderaa, R. B. (2000a). The Children's Social Behavior Questionnaire for milder variants of PDD problems: Evaluation of the psychometric characteristics. *Journal of Autism and Developmental Disorders*, *30*, 317–330.

Luteijn, E. F., Serra, M., Jackson, A. E., Steenhuis, M. P., Althaus, M., Volkmar, F. R., & Minderaa, R. B. (2000b). How unspecified are disorders of children with a Pervasive Developmental Disorder Not otherwise Specified? A study of social problems in children with PDDNOS and or ADHD. *European Child and Adolescent psychiatry*, *9*, 168–179.

Mahoney, W. J., Szatmari, P., MacLean, J. E., Bryson, S. E., Bartolucci, G., Walter, S. D., Jones, M. B., & Zwaigenbaum, L. (1998). Reliability and accuracy of differentiating pervasive developmental disorder subtypes. *Journal of the American Academy of Child and Adolescent Psychiatry*, *37*, 278–285.

Mayes, L. C., Volkmar, F. R., Hooks, M., & Cicchetti, D. (1993). Differentiating pervasive developmental disorder not otherwise specified from autism and language disorders. *Journal of Autism and Developmental Disorders*, *23*, 79–90.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.

Perry, R. (1998). Misdiagnosed ADD/ADHD; rediagnosed PDD. *Journal of the American Academy of Child and Adolescent Psychiatry*, *37*, 113–114.

Prior, M., Leekam, S., Ong, B., Eisenmajer, R., Wing, L., Gould, J., & Dowe, D. (1998). Are there subgroups within the autistic spectrum? A cluster analysis of a group of children with autistic spectrum disorders. *Journal of Child Psychology and Psychiatry*, *39*, 893–902.

Rutter, M., Bailey, A., & Lord, C. (2003). *Social Communication Questionnaire*. Los Angeles: Western Psychological Services.

Rutter, M., Caspi, A., & Moffit, T. E. (2003). Using sex differences in psychopathology to study causal mechanisms: Unifying issues and research strategies. *Journal of Child Psychology and Psychiatry*, *44*, 1092–1115.

Shaffer, D., Fisher Lucas, P. C. P., Dulcan, M. K., & Schwab-Stone, M. E. (2000). NIMH Diagnostic Interview Schedule of Children Version-IV (NIMH DISC IV): Description, differences from previous versions, and reliability of some common diagnoses. *Journal of the American Academy of Child and Adolescent Psychiatry*, *39*, 28–38.

Smalley, S. L. (1997). Genetic influences in childhood-onset psychiatric disorders: Autism and attention-deficit/hyperactivity disorder. *American Journal of Human Genetics*, *60*, 1276–1282.

Sparrow, S. S., Balla, D., & Cicchetti, D.V. (1984). *Vineland adaptive behavior scales* (survey ed.). Circle Pines, MN: American Guidance Service.

Steffenburg, S., & Gillberg, C. (1986). Autism and autistic-like conditions in Swedish rural and urban areas: A population study. *The British Journal of Psychiatry*, *149*, 81–87.

Stevens, J. P. (2002). *Applied multivariate statistics for the social sciences* (4th ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Towbin, K. E. (1997). Pervasive developmental disorders not otherwise specified. In: D. J. Cohen, & F. R. Volkmar (Eds.), *Handbook of autism and pervasive developmental disorders*. (pp. 123–147). New York: John Wiley & Sons.

Verhulst, F. C., van der Ende, J., & Koot, J. M. (1996). *Handleiding voor de CBCL/4–18*. Rotterdam: Afdeling kinder- en jeugdpsychiatrie, Sophia kinderziekenhuis/Academisch Ziekenhuis Rotterdam/Erasmus Universiteit Rotterdam.

Volkmar, F. R., Klin, A., & Cohen, & F. R. Volkmar (Ed.), *Handbook of autism and developmental disorders*. (pp. 5–40). New York: John Wiley & Sons.

Walker, D. R., Thompson, A., Zwaigenbaum, L., Goldberg, J., Bryson, S. E., Mahoney, W. J., Strawbridge, C. P., & Szatmari, P. (2004). Specifying pdd-nos: A comparison of pdd-nos, asperger syndrome, and autism. *Journal of the American Academy of Child and Adolescent Psychiatry*, *43*, 172–180.

Waterhouse, L., Morris, R., Allen, D., Dunn, M., Fein, D., Feinstein, C., Rapin, I., & Wing, L. (1996). Diagnosis and classification in autism. *Journal of Autism and Developmental Disorders*, *26*, 59–86.

World Health Organisation. (1994). *The ICD-10 classification of mental and behavioural disorders. Clinical descriptions and guidelines*. Geneva: WHO.

Wing, L. (1981). Language, social, and cognitive impairments in autism and severe mental retardation. *Journal of Autism and Developmental Disorders*, *11*, 31–44.

Wing, L., & Gould, J. (1979). Severe impairments of social interaction and associated abnormalities in children: Epidemiology and classification. *Journal of Autism and Developmental Disorders*, *9*, 11–29.