Elsevier Editorial System(tm) for Research in Autism Spectrum Disorders

Manuscript Draft

Manuscript Number: RASD-D-19-00191R1

Title: Early Screening for Autism Spectrum Disorder in Young Children with Developmental Problems Using the Chinese Version of the Child Behavior Checklist

Article Type: Research Paper

Keywords: autism spectrum disorder; Child Behavior Checklist/1.5-5 (CBCL/1.5-5); emotional and behavioral problems; sensitivity; specificity

Corresponding Author: Professor Chin-Chin Wu,

Corresponding Author's Institution:

First Author: Lai-Sang Iao

Order of Authors: Lai-Sang Iao; Wen-Han Yu; Chin-Chin Wu

Abstract: Background: Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder characterized by impairments in social interaction and communication, as well as repetitive patterns of behaviors and restricted interests. Current evidence suggested that children with ASD exhibited high level of co-occurring behavioral and emotional problems. Thus, this study was aimed to investigate the utility of the Chinese version of the Child Behavior Checklist for Ages 1.5-5 (CBCL/1.5-5) for assessing and screening ASD at an early age in Taiwan.
Methods: The CBCL/1.5-5 was completed by the caregivers of 134 clinically referred young children aged 18-47 months, including 66 children with ASD and 68 children with developmental delay (DD).
Results: The findings of this study showed that young children with ASD scored significantly higher than young children with DD for the following scales: Internalizing, Anxious/Depressed, Withdrawn, Attention Problems and Pervasive Developmental Problems (PDP). The results also showed that the Withdrawn scale yielded the best discrimination between the two groups using a T-score of 66 as cutoff. The area under the curve, sensitivity and specificity were .83, .74, and .77 respectively.
Conclusion: The findings of this study supported that the CBCL/1.5-5, especially the Withdrawn scale of the syndrome scale, could be used to differentiate young children with ASD from those with DD in Taiwan. Replication with a larger sample size is needed to validate the findings.

Early Screening for Autism Spectrum Disorder in Young Children with Developmental Problems Using the Chinese Version of the Child Behavior Checklist

Lai-Sang Iao[a] Wen-Han Yu[b] Chin-Chin Wu[c]*

[a]Department of Psychology, Nottingham Trent University, Nottingham, United Kingdom

[b]Jianan Psychiatric Center, Ministry of Health and Welfare, Taiwan

[c]Kaohsiung Medical University; Kaohsiung Medical University Hospital, Taiwan

Chin-Chin Wu (Corresponding author), Kaohsiung Medical University; Kaohsiung Medical University Hospital, e-mai:jinnchin@mail2000.com.tw

Conflict of interest: The authors declare that they have no conflict of interest.

Early Screening for Autism Spectrum Disorder in Young Children with Developmental Problems Using the Chinese Version of the Child Behavior Checklist

Abstract

Background: Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder characterized by impairments in social interaction and communication, as well as repetitive patterns of behaviors and restricted interests. Current evidence suggested that children with ASD exhibited high level of co-occurring behavioral and emotional problems. Thus, this study was aimed to investigate the utility of the Chinese version of the Child Behavior Checklist for Ages 1.5-5 (CBCL/1.5-5) for assessing and screening ASD at an early age in Taiwan.

Methods: The CBCL/1.5-5 was completed by the caregivers of 134 clinically referred young children aged 18-47 months, including 66 children with ASD and 68 children with developmental delay (DD).

Results: The findings of this study showed that young children with ASD scored significantly higher than young children with DD for the following scales: Internalizing, Anxious/Depressed, Withdrawn, Attention Problems and Pervasive Developmental Problems (PDP). The results also showed that the Withdrawn scale yielded the best discrimination between the two groups using a T-score of 66 as cutoff. The area under the curve, sensitivity and specificity were .83, .74, and .77 respectively.

Conclusion: The findings of this study supported that the CBCL/1.5-5, especially the Withdrawn scale of the syndrome scale, could be used to differentiate young children with ASD from those with DD in Taiwan. Replication with a larger sample size is needed to validate the findings.

Key words: autism spectrum disorder, Child Behavior Checklist/1.5-5 (CBCL/1.5-5), emotional and behavioral problems, sensitivity, specificity

Introduction

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder characterized by impairments in social interaction and communication, as well as repetitive patterns of behaviors and restricted interests with an estimated prevalence of 1% according to the 5th edition of *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5; American Psychiatric Association [APA], 2013). Recent studies show that a diagnosis of ASD could be made before 36 months with high stability over time (Ozonoff et al., 2015; Woolfenden, Sarkozy, Ridley, & Williams, 2012) and early intervention contributes to better outcomes and lower long-term cost of care (Peters-Scheffer, Didden, Korzilius, & Matson, 2012; Strauss et al., 2012). However, early diagnosis of ASD is still a challenge across countries, mainly due to considerable variations in symptoms and severity, insufficient expertise with extensive training and lengthy assessment for each child (Durkin et al., 2015; Limberg, Gruber, & Noterdaeme, 2017; Wong et al., 2018). Using standardized parent-report checklists that can efficiently and effectively screen for ASD for further assessment is thus an essential strategy for early identification. Among a variety of standardized parent-report checklists that are specifically developed for screening ASD (for reviews, see Barton, Dumont-Mathieu, & Fein, 2012), the Modified Checklist for Autism in Toddlers (M-CHAT; Robins, Fein, Barton, & Green, 2001) has been translated and validated for use in East Asia, including Taiwan (Wong et al., 2018), South Korea (Seung et al., 2015) and Japan (Kamio et al., 2014). However, in Taiwan, one would only administer the M-CHAT when one recognizes the possibility of ASD. Moreover, the scores of the M-CHAT could not provide robust information for making a diagnosis of ASD. We need more information (e.g., parents' concerns, child's

behavioral and emotional problems) for making formal clinical diagnosis given that behavioral and emotional problems are frequently found in children with ASD (e.g., Gau et al., 2010; Gray et al., 2012) and they have a crucial impact on parental stress and early intervention (e.g., Giovagnoli et al., 2015; Hou, Stewart, Iao, & Wu, 2018). A broadband behavior rating scale that provides a profile of behavioral and emotional problems is thus likely to be used in addition to the M-CHAT. If this broadband behavior rating scale also serves well as an early screener for ASD, it would help in improving early detection of ASD, particularly in Taiwan as the M-CHAT is the only standardized parent-report checklist for screening ASD that is available in clinical settings. This study thus examined the utility of a widely used broadband behavior rating scale for screening ASD at an early age in clinical settings in Taiwan.

The Child Behavior Checklist (CBCL; Achenbach & Rescorla, 2000) is a widely used broadband behavior rating scale that has been increasingly used and reported as a screener for ASD across countries since the 1980s (Havdahl, Tetzchner, Huerta, Lord, & Bishop, 2016; Limberg et al., 2017; Myers, Gross, & McReynolds, 2014; Muratori et al., 2011; Rescorla, 1988; Rescorla, Kim, & Oh, 2015; Sikora, Hall, Hartley, Gerrard-Morris, & Cagle, 2008). It is also the only broadband behavior rating scale that is widely used in Taiwan. One advantage of it is that it provides a profile of behavioral and emotional problems that is particularly informative for the development of intervention plan. Another advantage of it relates to the fact that it assesses a wide range of behavioral and emotional problems rather than ASD in specific so parents' responses are less likely to be biased depending on whether they believe their child has an ASD. It has two versions, one for young children aged 1.5 to 5

years, and the other for children and adolescents aged 6 to 18 years. This study focused on the one for young children aged 1.5 to 5 years (i.e., the CBCL/1.5-5) as an early screener for ASD.

The CBCL/1.5-5 has 99 items that are scored on seven syndrome scales, five DSM-oriented scales, and three broadband scales (Achenbach & Rescorla, 2000). The seven syndrome scales include Emotionally Reactive, Anxious/Depressed, Somatic Complaints, Withdrawn, Attention Problems, Aggressive Behavior and Sleep Problems. All seven syndrome scales constitute one of the broadband scales, (i.e., the Total Problems scale). The other broadband scales are the Internalizing scale (i.e., the Emotionally Reactive, Anxious/Depressed, Somatic Complaints, and Withdrawn scales of the syndrome scales) and the Externalizing scale (i.e., the Attention Problems and Aggressive Behavior scales of the syndrome scales). The five DSM-oriented scales are Affective Problems, Anxiety Problems, Attention Deficit/Hyperactivity Problems, Oppositional Defiant Problems, and Pervasive Developmental Problems (PDP). The PDP scale was renamed the Autism Spectrum Problems scale, and one item (i.e., *item 3. Afraid to try new things*) was removed to match with the DSM-5 diagnostic category of ASD in the latest version of the CBCL/1.5-5 (Achenbach, 2014). Out of all the scales, the Withdrawn scale of the syndrome scales and the former PDP scale of the DSM-oriented scales have been suggested as the best discrimination of children with and without ASD by several studies across countries (e.g. Limberg et al., 2017; Myers et al., 2014; Rescorla et al., 2015).

Sikora et al. (2008) is one of the first studies using the CBCL/1.5-5 to identify ASD in 147 American children aged 3-5 years who have demonstrated developmental problems. Using the cutoff of ≥ 70, the Withdrawn scale had a sensitivity (i.e., correctly identify a child who might have ASD)

of .65 and a specificity (i.e., correctly identify a child who does not have ASD) of .62. The PDP scale had a sensitivity of .80 and a specificity of .42. Myers et al. (2014) discriminated ASD from other developmental or behavioral problems in an American sample of 156 clinically referred children aged 2-5 years. Receiver operative characteristic (ROC) analyses revealed an area under the curve (AUC) of .75 for the Withdrawn scale and an AUC of .71 for the PDP scale. Using a cutoff of ≥ 65, the Withdrawn scale had a sensitivity of .89 and a specificity of .52 while the PDP scale had a sensitivity of .93 and a specificity of .29. More recently, Levy et al. (2019) discriminated children with ASD from children with developmental delay (DD) without ASD features, children with DD plus ASD features and controls in a large sample of 2413 American 3- to 5-year-olds. Using a cutoff of ≥ 65, the Withdrawn scale had a sensitivity of .66 and the PDP scale had a sensitivity of .80 regardless of the comparison group. Specificities of the Withdrawn scale were .96 for the ASD-control comparison, .91 for the ASD-DD comparison, and .63 for the ASD-DD+ASD comparison whereas the specificities of the PDP scale were .93, .85, and .50 respectively.

Apart from the United States, the CBCL/1.5-5 has also been investigated in European. Muratori et al. (2011) identified ASD from typical development and other psychiatric disorders (OPD) in 313 Italian children aged 2-5 years. Using a cutoff of ≥ 65, they reported a sensitivity of .89 for the Withdrawn scale and .85 for the PDP scale when comparing children with ASD against children with typical development and children with OPD. However, AUCs and specificities varied depending on the group to which children with ASD were compared. When children with ASD were compared to children with typical development, AUC was .95 for both the Withdrawn scale and the PDP scale, and

specificity were .92 and .90 respectively. When the ASD group was compared to the OPD group, AUCs dropped to .85 for the Withdrawn scale and .81 for the PDP scale, and specificities dropped to .65 and .60 respectively. More recently, Limberg et al. (2017) distinguished ASD from OPD in 183 German children aged 2-5 years. With a cutoff of ≥ 60.5 on the Withdrawn scale, AUC was .81, sensitivity was .88 and specificity was .63. With a cutoff of ≥ 64.5 on the PDP scale, AUC was .78, sensitivity was .83 and specificity was .60.

Rescorla et al. (2015) is the only study, so far, that examined the CBCL/1.5-5 in a non-Western country. They screened ASD from DD, OPD and typical development in Korean children with a mean age of 3.4 years. Using a cutoff of ≥ 65, sensitivity was .78 for the Withdrawn scale and .80 for the PDP scale when the ASD group was compared to the other three groups. Similar to Muratori et al. (2011), AUCs and specificities varied depending on the group to which children with ASD were compared. When compared to children with typical development, AUC was .94 for the Withdrawn scale and .93 for the PDP scale, and specificities were .89 and .87 respectively. When compared to children with OPD, AUC were .74 and .70, and specificities were .63 and .55 respectively. When compared to children with DD, AUC were .73 and .68, and specificities were .53 and .60 respectively.

According to Cicchetti, Volkmar, Klin, and Showalter (1995), diagnostic accuracy suggested by AUC, sensitivity and specificity are poor (≤ .70), fair (.70-.79), good (.80-.89), and excellent (> .90). Therefore, previous studies suggested that both the Withdrawn scale and the PDP scale showed good to excellent diagnostic accuracy in identifying children with ASD from children with typical development, implying that both scales are suitable for Level 1 screening of ASD in general

populations. However, specificities for both scales fell to poor when discriminating children with ASD from clinically referred children (i.e., children with OPD and children with DD). This suggests that the two scales may be less suitable for Level 2 screening of ASD in clinically referring populations as they may falsely identify clinically referred children without ASD as at risk of ASD. These children will be referred for time-consuming assessments using the Autism Diagnostic Interview-Revised (ADI-R; Rutter, Conteur, & Lord, 2003) and/or the Autism Diagnostic Observation Schedule (ADOS; Lord, Rutter, DiLavore, & Risi, 1999) that may not be necessary. Although this seems to be a global challenge in diagnosing ASD, it has a greater impact on low- and middle-income settings where healthcare resources are limited.

Taiwan is one of these low- and middle-income settings where an accurate differentiation of children with ASD from clinically referred children at an early age is essential for restricted resources to be effectively targeted at those who are in need. As mentioned earlier, the M-CHAT is the only standardized Level 2 screener for ASD in Taiwan (Wong et al., 2018) and it is usually used alongside with the CBCL/1.5-5. If the CBCL/1.5-5 could also serve well as an early Level 2 screener for ASD, using both the M-CHAT and the CBCL/1.5-5 would help in improving early and accurate detection of ASD in Taiwan. However, it is not clear whether the CBCL/1.5-5 could be used as an early Level 2 ASD screener in Taiwan. Previous studies have showed that preschoolers in Taiwan scored higher on the CBCL/1.5-5 than their counterparts in America (Wu et al., 2012) and in 23 societies (Rescorla et al., 2011). It is thus possible that cutoffs would need to be higher when using the CBCL/1.5-5 as an early Level 2 ASD screener in Taiwan. Moreover, Stewart and Lee (2017) suggested that establishing

culturally optimal cutoffs is essential for populations in which there are no histories of screening ASD with a particular instrument. This study thus aimed to identify the significant scales of the CBCL/1.5-5 and determine the optimal cutoffs for use in Taiwan that would best differentiate young children with ASD from young children with DD in clinical setting.

One point to note from the previous studies is that none of them matched the different groups of children with chronological age, mental age or intelligence (i.e., IQ). These factors are widely known to have confounding effects on findings if groups for comparisons differ on these factors in addition to target variables. Narzisi et al. (2013) is one of very few studies that matched three groups of 141 Italian children aged 1.5-3 years by chronological age, gender, race and socioeconomic status. They found good to excellent AUCs, sensitivities and specificities (.94, .90 and .83 respectively for the Withdrawn scale with a cutoff of $\geq$ 65, and .91, .85 and .83 respectively for the PDP scale with a cutoff of $\geq$ 69) when ASD was compared with OPD. When ASD was compared with typical development, very high AUCs, sensitivities and specificities of > .90 for both scales with a cutoff of $\geq$ 65 were obtained. As a result, their findings suggested a higher diagnostic accuracy of the two scales than the previous studies and this may be due to the matching procedure. In addition to chronological age, gender, race and socioeconomic status, mental age should also be matched between groups of children given the high incidence of low IQs among children with ASD (Hermelin & O'Connor, 1970). With all of these factors matched in this study, we expected that children with ASD would show more behavioral and emotional problems than children without ASD in a clinically referred sample. We also expected that both the Withdrawn scale and the PDP scale would demonstrate high

diagnostic accuracy in supporting a precise and early ASD diagnosis when one recognizes developmental and behavioral problems in a young child in Taiwan. Note that the PDP scale (Achenbach & Rescorla, 2000) instead of the latest Autism Spectrum Problems scale (Achenbach, 2014) was used so that a comparison could be made with previous studies.

Method

Participants

Young children were suspected of having problems with development were referred from an early intervention center and the department of psychiatry of a teaching hospital in the Southwest area of Taiwan. At the time of being investigated, only one child was under 18 months (i.e., 16 months), and the rest were between the ages of 18 and 47 months. Informed consent was obtained from 134 children's parents prior to assessment. All participants (103 boys, 31 girls) were assessed and diagnosed by a multidisciplinary team that included senior child psychologists with Ph.D. degree and child psychiatrists. All children failed to reach a total score of 85 on the Mullen Scales of Early Learning (MSEL; Mullen, 1995) or a T-score of 35 on any of the four cognitive scales (i.e., visual reception, fine motor, receptive language, and expressive language), suggesting they had DD. Sixty-six of these young children were later diagnosed with ASD according to Frazier et al.'s (2012) suggestion on the DSM-5 criteria for ASD[1] and with reference to developmental history, daily activities

[1] Frazier et al. (2012) suggested a relaxed algorithm, requiring one less deficit in either social communication/interaction or restricted/repetitive behaviors in order to increase the sensitivity of the

and performance, parental concerns, clinical observation, and the results of the ADOS (Lord et al., 1999). Of these 66 children, 50 (76%) children met the strict DSM-5 criteria for ASD. The rest of the young children (i.e., 68) were diagnosed with DD only with reference to developmental history, adaptive behavior, and the results of the MSEL. However, it is important to note that there were only 5 (7%) of the 68 children were found without any ASD features (i.e., their ADOS total scores were 0). As a result, most children in the DD group showed some ASD features (i.e., mean of their ADOS total scores were 3.40) but not to a clinically significant level. Participant characteristics are presented in Table 1. The two groups of children did not differ in chronological age, verbal mental age, nonverbal mental age, overall mental age, social-economic status and ratio of gender. However, the two groups differed on the ADOS score, with the ASD group showing higher symptom severity.

Insert Table 1

Materials and Procedure

Participants were accompanied by their parents and were individually assessed with the MSEL (Mullen, 1995) and the ADOS (Lord et al., 1999). The MSEL was administered by well-trained graduate students studying in a Master of Science program in clinical psychology under the supervision of a senior child psychologist with Ph.D. degree whereas the ADOS was administered by

---

DSM-5 for ASD. According to the DSM-5, children would meet the criteria for an ASD diagnosis if they manifest (1) three deficits in social communication/interaction and one restricted/repetitive behavior, or (2) two deficits in social communication/interaction and two restricted/repetitive behaviors.

a senior child psychologist with Ph.D. degree. The MSEL is a standardized developmental test for assessing cognitive and language skills in children from birth to 68 months. T-scores and age equivalents were computed for each of the four cognitive scales. These age equivalents were then summed together and divided by four, constituting a child's overall mental age. Verbal mental age was determined by averaging the age equivalents obtained from the expressive language and receptive language subscales of the MSEL. The same was done for the nonverbal mental age using the visual reception and fine motor subscales of the MSEL. The MSEL manual reports acceptable reliability in internal consistency, retesting and interrating. It has also demonstrated concurrent validity with other well-known developmental tests of language and cognitive development such as Bayley Scales of Infant Development (BSID; Bayley, 1969).

The ADOS (Lord et al., 1999) is a semi-structured, play-based and observational assessment for individuals who are suspected of having ASD. It provides a standardized context for observing and scoring communication, social and stereotyped behaviors, and restricted interests. It consists of four modules, each was designed for individuals of different levels of expressive language. All participants were administered with Module 1 and it took 30 to 45 minutes. Diagnosis was made on the basis of the three domain scores: social interaction, communication and a combined social-communication total. A higher score reflects more severe autistic symptoms.

While participants were being assessed, parents were asked to complete the CBCL/1.5-5 which is a widely used questionnaire for measuring behavioral and emotional problems in children aged 1.5 to 5 years (Achenbach & Rescorla, 2000). It contains 99 items. Each item is rated 0 for not

true, 1 for somewhat or sometimes true, and 2 for very true or often true, based on the display of behavioral and emotional problems in a child from the last 2 months. Ratings are used to compute T scores for the seven syndrome scales, five DSM-oriented scales, and three broadband scales using the computer scoring program with US norms. For both seven syndrome scales and five DSM-oriented scales, the cutoff for normal range is a T score < 65, borderline is from 65 to 69, and the clinical range is ≥ 70. For three broadband scales, the cutoff for normal range is a T score < 60, borderline is from 60 to 63, and the clinical range is ≥ 64. The test-retest reliability and construct validity of the CBCL/1.5-5 in a preschool sample in Taiwan were 0.52-0.84 and 0.37-0.91 respectively (Wu et al., 2012). The whole session lasted approximately 2 hours.

Data Analysis

Children's T scores for each CBCL/1.5-5 scale was first examined. Chi-square analysis with Fisher's Exact Test (Fisher, 1935) correction was used to test whether there were more young children with ASD than young children with DD who reached the borderline range of 65 to 69 and the clinical range of ≥ 70 for each of the scales. Independent t-tests were also performed to test whether young children with ASD scored higher than young children with DD for each of the scales. An alpha level of $p < .003$ was used due to the large number of comparisons. A discriminant analysis was then conducted with the scales that significantly discriminated the two groups at $ps < .01$. These scales were ranked according to their loadings on the discriminant function using the structure matrix in the discriminant analysis. Finally, ROC analyses were performed to determine the optimal cutoffs for the scales that were identified in the discriminant analysis. To further evaluate the diagnostic accuracy of

each identified scale, AUC was also examined.

Results

Table 2 presents the CBCL/1.5-5 means and standard deviations for the two groups of children. Independent t-tests showed that young children with ASD scored significantly higher than young children with DD for the following scales: Internalizing, Anxious/Depressed, Withdrawn, Attention Problems and PDP. Table 3 shows the percentages of children who have reached the borderline range and above (including the clinical range) for each of the subscales. There were significantly more young children with ASD than young children with DD who met the borderline and clinical range for the following scales: Withdrawn, Attention Problems, PDP and Total Problems.

Insert Table 2 and 3

Following the results of the independent t-tests, the scales with *ps* < .003 (e.g., Internalizing, Withdrawn) were further examined in the discriminant analysis. The results (Wilks' Lambda = .64, *p* < .001) suggested that the discriminant function could be reliably used to distinguish young children with ASD from those with DD. The group centroids were .75 for ASD and -.73 for DD. For a total of 99 children (73.9%) that were classified accurately, 44 were young children with ASD (66.7%) and 55 were young children with DD (80.2%). Table 4 reports the structure matrix showing the simple Pearson correlation of each scale with the discriminant function. These correlations are called structure coefficients or discriminant loadings. The larger the coefficient or loading is, the greater the

contribution of the respective scale to the discrimination between the two groups is. As a result, the Withdrawn scale is the key scale in discriminating the two groups, followed in decreasing order by the PDP scale and the Internalizing scale.

Insert Table 4

In the ROC analyses, the optimal range of cutoffs for the Withdrawn scale, the PDP scale and the Internalizing scale were 60 to 70, 67 to 71, and 60 to 62 respectively. Table 5 presents the sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) for each cutoff for the Withdrawn, the PDP, and the Internalizing scales. In addition, the AUCs of the three scales were .83, .77, and .70 respectively.

Insert Table 5

We further examined the means of the items in the Withdrawn scale and the PDP scale for the two groups (Table 6). Independent t-tests showed that young children with ASD scored significantly higher than young children with DD for all items except Item 3, 7, 21, 63, 76, 92 from the PDP scale. In addition, we examined the percentage of children in both groups who rated 0 and either 1 or 2 on the items from the Withdrawn scale and the PDP scale. The findings showed that there were significantly more young children with ASD than young children with DD who rated either 1 or 2 on the

shared Item 4, 67, 98 from both the Withdrawn and the PDP scale and all of the other items from the Withdrawn scale and Item 25 and 80 from the PDP scale (Table 7).

Insert Table 6 and 7

Discussion

This study examined the diagnostic utility of the CBCL/1.5-5 for screening ASD at an early age in clinical settings in Taiwan. Young children with ASD and young children with DD were well-matched by a variety of demographic characteristics, including chronological age, verbal mental age, nonverbal mental age, overall mental age, social-economic status and ratio of gender. The results showed that young children with ASD scored significantly higher than young children with DD on the following scales of the CBCL/1.5-5: Internalizing, Anxious/Depressed, Withdrawn, Attention Problems and PDP. Their scores were not significantly different on the other scales of the CBCL/1.5-5. These findings supported the hypothesis that more behavioral and emotional problems were presented in children with ASD than children without ASD in a clinically referred sample. In particular, the Withdrawn scale of the syndrome scale, the PDP scale of the DSM-oriented scale and the Internalizing scale of the broadband scale were the three scales that showed largest differences between the two groups of ASD and DD. Consistent to previous studies (e.g., Limberg et al., 2017; Myers et al., 2014; Rescorla et al., 2015), the Withdrawn scale of the syndrome scale and the PDP scale of the DSM-oriented

scale demonstrated the best discrimination between the two groups suggested by a fair to good AUC. Therefore, our second hypothesis was also supported. Findings of this study and their implications are further discussed below.

Our literature review suggested that the CBCL/1.5-5 scales, more specifically the Withdrawn scale and the PDP scale, were better at discriminating children with ASD from typically developing children as Level 1 screening tool than clinically referred children as Level 2 screening tool based on the low specificities reported in previous studies (Limberg et al., 2017; Muratori et al., 2011; Myers et al., 2014; Narzisi et al., 2013; Rescorla et al., 2015). These low specificities were yielded when comparing children with ASD against clinically referred children who were not matched by demographic characteristics such as chronological age and mental age. When the comparison groups were matched on various demographic characteristics, Narzisi et al. (2013) and the current study indicated higher specificities for the Withdrawn scale (i.e., ≥ 77% with a cutoff of ≥ 64) and the PDP scale (i.e., ≥ 65% with a cutoff of ≥ 69). However, these higher specificities may also be due to the young age of the participants in Narzisi et al.'s study and the current study. One of the explanations that Narzisi et al. proposed for their findings was that parents of older children might be more accustomed to the atypical behaviors of their children than parents of younger children. This could also be the case for the current study.

It is also important to note that higher specificities are inevitably at the cost of lower sensitivities. It would be worse for an ASD screener to miss children with ASD (i.e., low sensitivity) than falsely identified children without ASD as at risk of ASD (i.e., low specificity). This is because missing

children with ASD would lead to serious and long-term consequences to their families and community whereas falsely identifying children without ASD as at risk of ASD would only lead to further assessment that may not be necessary. Using a cutoff of 60 for the Withdrawn scale and 67 for the PDP scale would include more children as having ASD especially that our sample was a high-risk group. Nevertheless, one should also consider that higher sensitivity and lower specificity would increase worry and anxiety in parents of children with developmental problems as previous studies suggested that stigmatization exists in Chinese culture (Pang et al., 2018). The current study therefore suggests a cutoff of 66 for the Withdrawn scale and 70 for the PDP scale based on the diagnostic accuracy indicated by sensitivity, specificity, PPV, NPV as well as AUC in Table 5. These cutoffs are slightly higher than those used in previous studies (e.g., Narzisi et al., 2013; Rescorla et al., 2015; Levy et al., 2019). This difference may relate to the fact that the current study matched the groups on various demographic characteristics, including mental age, while previous studies did not. To accurately discriminate children with DD who also had ASD from those with DD, slightly higher cutoffs may be needed. Another possible explanation could be related to cultural differences. Preschoolers in Taiwan tend to score higher on the CBCL/1.5-5 than their counterparts in other societies (Rescorla et al., 2011; Wu et al., 2012). The same trend was also observed in most of the CBCL/1.5-5 scales when comparing the T scores of the current study with those reported in Giovagnoli et al. (2015) and Muratori et al. (2011), which further suggested that identifying culturally optimal cutoffs for the Taiwan population is necessary.

The results of this study indicated that the diagnostic accuracy was fair to good when using 66

as the cutoff for the Withdrawn scale. However, it was poor to fair using 70 as the cutoff for the PDP scale. Table 6 also showed that all items in the Withdrawn scale significantly discriminated the two groups while there were only 7 out of the 13 items in the PDP scale did so. This difference between the two scales did not only demonstrate in mean score for each item (Table 6) but also in number of participants who scored more than 0 for each item (Table 7). Hence, the Withdrawn scale yielded better decision statistics and provided a better discrimination between the two groups than the PDP scale. Moreover, the Withdrawn scale did better with a lower cutoff than the PDP scale. These findings were consistent with earlier results reported in Table 2 and 3. In Table 2, the mean scores for the Withdrawn scale were slightly lower than those for the PDP scale and the effect size for the Withdrawn scale was larger than that of the PDP scale. In Table 3, the PDP scale seemed to have a higher false positive rate (51.5% of the DD group) than the Withdrawn scale (23.5%) when the cutoff for both scales was 65, suggesting a higher cutoff for the PDP scale compared to the Withdrawn scale was appropriate. This pattern of findings was also found in Sikora et al. (2008) and Rescorla et al. (2015).

One possible explanation for the different findings between the Withdrawn scale and the PDP scale, despite the fact that they share 5 identical items (e.g., *Item 4: Avoids looking others in the eye*; *Item 23: Doesn't answer when people talk to him/her*), could also be that the two groups in the current study were highly matched except that their ADOS scores were significantly different. The ASD group had DD and the DD group also showed some ASD features but not to a clinically significant level. Previous studies (e.g., Levy et al. 2019; Wiggins et al., 2015) suggested that children

with DD plus ASD features showed significantly more impairments on all cognitive, behavioral and social domains and were more similar to children with ASD compared to DD children without ASD features. Their findings also showed that it was not easy to distinguish children with ASD from children with DD plus ASD features when the cutoff for both Withdrawn and PDP scales was 65. Therefore, a higher cutoff may be necessary and our findings extended previous literature by suggesting that the Withdrawn scale may be better than the PDP scale in distinguishing children with ASD plus DD and children with DD plus ASD features.

Rescorla et al. (2015) also suggested that different clinically referred samples may lead to the difference found between the Withdrawn scale and the PDP scale. When children with ASD was compared to children with OPD, Rescorla et al. reported fair AUCs for both the Withdrawn scale (.74) and the PDP scale (.70). When children with ASD was compared to children with DD, they reported a fair AUC for the Withdrawn scale (.73) but a poor AUC for the PDP scale (.68). It is thus possible that the Withdrawn scale may be better than the PDP scale in discriminating children with ASD from children with DD but not children with OPD, which was the comparison group used in other previous studies (e.g., Narzisi et al., 2013). However, it is also possible that this finding is specific to East Asian culture. Given that the current study and Rescorla et al.'s study are the only studies that were conducted in East Asia, further investigation is needed.

In addition to the Withdrawn scale and the PDP scale, the Internalizing scale was ranked the third in differentiating children with ASD from children with DD in Taiwan. Although its sensitivities and specificities were mostly poor using the optimal cutoff range of 60 to 62, its AUC was fair. Previous

studies did not report any values of AUC, sensitivity and specificity for the Internalizing scale as their findings did not suggest it as significant in discriminating children with and without ASD. Muratori et al. (2011), Myers et al. (2014) and Rescorla et al. (2015) even did not find a significant difference between children with ASD and clinically referred children on their scores for the Internalizing scale. To our knowledge, only Narzisi et al. (2013) and Havdahl et al. (2016) indicated that children with ASD scored higher on the Internalizing scale than children with OPD. It is not clear why findings were not consistent for the Internalizing scale. As a result, the diagnostic utility of the Internalizing scale for screening ASD at an early age remains questionable.

As most of the previous studies (e.g. Limberg et al., 2017; Narzisi et al., 2013), one of the limitations of the current study is the small sample size that may limit the power of the study. Replication with larger sample size which may better reflect the high heterogeneity of ASD is needed to validate the findings. Future studies should also aim to recruit participants from an epidemiological sample to further test the generalizability of the findings and the diagnostic utility of the CBCL/1.5-5 as a Level 1 screener of ASD in Taiwan. Another limitation is that the ADOS (Lord et al., 1999) used in the assessment procedure may not be as reliable as the second edition of the ADOS (Lord, Rutter, DiLavore, Risi, Gotham, & Bishop, 2012), which involves a toddler module, in assessing young children (Luyster et al., 2009). However, each participant in this study was diagnosed by a multidisciplinary team that included senior child psychologists and experienced child psychiatrists with reference to the child's developmental history, daily activities and performance, parental concerns, and clinical observation. Misdiagnosis was thus very unlikely.

To conclude, our findings suggested that the CBCL/1.5-5, especially the Withdrawn scale of the syndrome scale, can be used to screen ASD at an early age in a clinically referred sample of children with fairly good diagnostic accuracy in Taiwan. When one recognizes developmental and behavioral problems in a young child, the CBCL/1.5-5 can be efficiently completed to not only support an early diagnosis of ASD by a multidisciplinary team in clinical settings but also provide a profile of behavioral and emotional problems for an early intervention plan in various settings in Taiwan. With further investigation for generalizability, the Chinese version of the CBCL/1.5-5 could be a useful instrument for screening ASD at an early age in East Asia where Chinese populations are prevalent to effectively target restricted resources for assessment and intervention at high risk children so that better developmental outcomes and lower long-term cost of familial and community care could be yielded.

Compliance with Ethical Standards

Funding

This study was supported by the Ministry of Science and Technology (MOST-103-2628-H-037-001-MY2, MOST-105-2410-H-037-001-MY3).

Ethical Approval

All procedures performed in this study involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. This study was approved by the Ditmanson Medical Foundation Chia-Yi Christian Hospital Research Ethics Committee (CYCH-IRB102045).

Informed consent

Informed consent was obtained from all individual participants included in this study.

Conflict of interest

The authors declare that there are no conflicts of interest.

References

Achenbach, T. M. (2014). *DSM-oriented guide for the Achenbach System of Empirically Based Assessment (ASEBA)*. Burlington, VT: Vermont Center for Children, Youth and Families, University of Vermont.

Achenbach, T. M., & Rescorla, L. A. (2000). *Manual for the ASEBA preschool forms and profiles*. Burlington, VT: Research Center for Children, Youth, and Families, University of Vermont.

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: American Psychiatric Association.

Barton, M. L., Dumont-Mathieu, T., & Fein, D. (2012). Screening young children for autism spectrum disorders in primary practice. *Journal of Autism and Developmental Disorders, 42*(6), 1165-1174.

Bayley, N. (1969). *Bayley Scales of Infant Development*. San Antonio, TX: Psychological Corporation.

Cicchetti, D. V., Volkmar, F., Klin, A., & Showalter, D. (1995). Diagnosing autism using ICD-10 criteria: A comparison of neural networks and standard multivariate procedures. *Child Neuropsychology, 1*, 26-37.

Durkin, M. S., Elsabbagh, M., Barbaro, J., Gladstone, M., Happe, F., Hoekstra, R. A., ... Tager-Flusberg, H. (2015). Autism screening and diagnosis in low resource settings: Challenges and opportunities to enhance research and services worldwide. *Autism Research*, *8*(5), 473-476.

Fisher, R. A. (1935). *The design of experiments*. Oxford, England: Oliver & Boyd.

Frazier, T., Youngstrom, E., Speer, L., Embacher, R., Law., Constantino, J., …Eng, C. (2012). Validation of proposed DSM-5 criteria for autism spectrum disorder. *Journal of the American*

Academy of Child and Adolescent Psychiatry, 51(1), 28-40.

Gau, S. S. F., Chou, M. C., Lee, J. C., Wong, C. C., Chou, W. J., Chen, M. F., …Wu, Y. Y. (2010). Behavioral problems and parenting style among Taiwanese children with autism and their siblings. *Psychiatry and Clinical Neurosciences, 64*(1), 70-78.

Giovagnoli, G., Postorino, V., Fatta, L. M., Sanges, V., De Peppo, L., Vassena, L., …Mazzone, L. (2015). Behavioral and emotional profile and parental stress in preschool children with autism spectrum disorder. *Research in Developmental Disabilities, 45*, 411-421.

Gray, K., Keating, C., Taffe, J., Brereton, A., Einfeld, S., & Tonge, B. (2012). Trajectory of behavior and emotional problems in autism. *American Journal on Intellectual and Developmental Disabilities, 117*(2), 121-133.

Havdahl, K. A., von Tetzchner, S., Huerta, M., Lord, C., & Bishop, S. L. (2016). Utility of the child behavior checklist as a screener for autism spectrum disorder. *Autism Research, 9*(1), 33-42.

Hermelin, B., & O'Connor, N. (1970). *Psychological experiments with autistic children.* Oxford: Pergamon Press.

Hou, Y.-U., Stewart, L., Iao, L.-S., & Wu, C.-C. (2018). Parenting stress and depressive symptoms in Taiwanese mothers of young children with autism spectrum disorder: Association with children's behavioral problems. *Journal of Applied Research in Intellectual Disabilities,31*(6), 1113-1121.

Kamio, Y., Inada, N., Koyama, T., Inokuchi, E., Tsuchiya, K., & Kuroda, M. (2014). Effectiveness of using the Modified Checklist for Autism in Toddlers in two-stage screening of autism spectrum disorder at the 18-month health check-up in Japan. *Journal of Autism and Developmental*

Disorders, *44*(1), 194-203.

Levy, S. E., Rescorla, L. A., Chittams, J. L., Kral, T. J., Moody, E. J., Pandey, J., ... & Rosenberg, C. R. (2019). ASD Screening with the Child Behavior Checklist/1.5-5 in the Study to Explore Early Development. *Journal of Autism and Developmental Disorders*, 49(6), 2348-2357.

Limberg, K., Gruber, K., & Noterdaeme, M. (2017). The German version of the child behavior checklist 1.5-5 to identify children with a risk of autism spectrum disorder. *Autism, 21*(3), 368-374.

Lord, C. (1995). Follow-up of two-year-olds referred for possible autism. *Journal of Child Psychology and Psychiatry, 36*(8), 1365-1382.

Lord, C., Rutter, M., DiLavore, P. C., & Risi, S. (1999). *Autism diagnostic observation (ADOS)*. Los Angeles, CA: Western Psychological Services.

Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., & Bishop, S. L. (2012). *Autism diagnostic observation schedule*, second edition (ADOS-2) manual (Part I): modules 1-4. Torrance, CA: Western Psychological Services.

Luyster, R., Gotham, K., Guthrie, W., Coffing, M., Petrak, R., Pierce, K., …Lord, C. (2009). The Autism Diagnostic Observation Schedule-toddler module: A new module of a standardized diagnostic measure for autism spectrum disorders. *Journal of Autism and Developmental Disorders, 39*(9), 1305-1320.

Mullen, E. M. (1995). *Mullen Scales of Early Learning.* Circle Pines, MN: American Guidance Service.

Muratori, F., Narzisi, A., Tancredi, R., Cosenza, A., Calugi, S., Saviozzi, I., ...Calderoni, S. (2011).

The CBCL 1.5-5 and the identification of preschoolers with autism in Italy. *Epidemiology and Psychiatric Sciences, 20*(4), 329-338.

Myers, C. L., Gross, A. D., & McReynolds, B. M. (2014). Broadband behavior rating scales as screeners for autism? *Journal of Autism and Developmental Disorders, 44*(6), 1403-1413.

Narzisi, A., Calderoni, S., Maestro, S., Calugi, S., Mottes, E., & Muratori, F. (2013). Child behavior check list 1½–5 as a tool to identify toddlers with autism spectrum disorders: A case-control study. *Research in Developmental Disabilities, 34*(4), 1179-1189.

Ohta, M., Nagai, Y., Hara, H., & Sasaki, M. (1987). Parental perception of behavioral symptoms in Japanese autistic children. *Journal of Autism and Developmental Disorders, 17*(4), 549-563.

Ozonoff, S., Young, G. S., Landa, R. J., Brian, J., Bryson, S., Charman, T., ...Iosif, A. M. (2015). Diagnostic stability in young children at risk for autism spectrum disorder: A baby siblings research consortium study. *Journal of Child Psychology and Psychiatry, 56*(9), 988-998.

Pang, Y. L., Lee, C. M., Wright, M., Shen, J., Shen, B., & Bo, J. (2018). Challenges of case identification and diagnosis of autism spectrum disorders in China: A critical review of procedures, assessment, and diagnostic criteria. *Research in Autism Spectrum Disorders, 53,* 53–66.

Peters-Scheffer, N., Didden, R., Korzilius, H., & Matson, J. (2012). Cost comparison of early intensive behavioral intervention and treatment as usual for children with autism spectrum disorder in the Netherlands. *Research in Developmental Disabilities, 33*(6), 1763-1772.

Rescorla, L. (1988). Cluster analytic identification of autistic preschoolers. *Journal of Autism and Developmental Disorders, 18*(4), 475-492.

Rescorla, L. A., Achenbach, T. M., Ivanova, M. Y., Harder, V. S., Otten, L., Bilenberg, N., …Verhulst, F.C. (2011). International comparisons of behavioral and emotional problems in preschool children: parents' reports from 24 societies. *Journal of Clinical and Child Adolescent Psychology, 40*(3), 456-467.

Rescorla, L., Kim, Y. A., & Oh, K. J. (2015). Screening for ASD with the Korean CBCL/1.5-5. *Journal of Autism and Developmental Disorders, 45*(12), 4039-4050.

Robins, D. L., Fein, D., Barton, M., & Green, J. (2001). The modified checklist for autism in toddlers: A study investigating the early detection of autism and pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 31*(2), 131-144.

Rutter, M., Le Conteur, A., & Lord, C. (2003). *Autism Diagnostic Interview—Revised (ADI-R).* Los Angeles, CA: Western Psychological Services.

Seung, H., Ji, J., Kim, S. J., Sung, I., Youn, Y. A., Hong, G., ...Youm, H. K. (2015). Examination of the Korean modified checklist of autism in toddlers: Item response theory. *Journal of Autism and Developmental Disorders*, *45*, 2744-2757.

Sikora, D. M., Hall, T. A., Hartley, S. L., Gerrard-Morris, A. E., & Cagle, S. (2008). Does parent report of behavior differ across ADOS-G classifications: Analysis of scores from the CBCL and GARS. *Journal of Autism and Developmental Disorders, 38,* 440-448.

Stewart, L. A., & Lee, L. C. (2017). Screening for autism spectrum disorder in low-and middle-income countries: A systematic review. *Autism, 21*(5), 527-539.

Strauss, K., Vicari, S., Valeri, G., D'Elia, L., Arima, S., & Fava, L. (2012). Parent inclusion in early

intensive behavioral intervention: The influence of parental stress, parent treatment fidelity and parent-mediated generalization of behavior targets on child outcomes. *Research in Developmental Disabilities, 33*(2), 688-703.

Wiggins, L. D., Levy, S. E., Daniels, J., Schieve, L., Croen, L. A., DiGuiseppi, C., ... & Reynolds, A. (2015). Autism spectrum disorder symptoms among children enrolled in the Study to Explore Early Development (SEED). *Journal of Autism and Developmental Disorders, 45*(10), 3183-3194.

Woolfenden, S., Sarkozy, V., Ridley, G., & Williams, K. (2012). A systematic review of the diagnostic stability of autism spectrum disorder. *Research in Autism Spectrum Disorders, 6*(1), 345-354.

Wong, Y. S., Yang, C. C., Stewart, L., Chiang, C. H., Wu, C. C., & Iao, L. S. (2018). Use of the Chinese version modified checklist for autism in toddlers in a high-risk sample in Taiwan. *Research in Autism Spectrum Disorders*, *49*, 56-64.

Wu, Y. T., Chen, W. J., Hsieh, W. S., Chen, P. C., Liao, H. F., Su, Y. N., & Jeng, S. F. (2012). Maternal-reported behavioral and emotional problems in Taiwanese preschool children. *Research in Developmental Disabilities, 33*(3), 866-873.

Table 1

Demographic Characteristics of Sample by Diagnosis

	Group					
	ASD (*n* = 66)		DD (*n* = 68)			
Characteristic	Mean	*SD*	Mean	*SD*	t/χ^2	*p*
Chronological age (months)	31.49	7.30	29.71	7.89	1.35	.18
Nonverbal mental age (months)	24.42	6.99	25.02	6.13	-0.53	.60
Verbal mental age (months)	18.85	7.69	20.57	5.10	-1.54	.13
Overall mental age (months)	21.64	7.00	22.83	5.29	-1.12	.27
Socioeconomic status	62.68	16.81	59.71	17.29	1.01	.31
ADOS score	16.29	3.18	3.15	1.76	29.72	.00*
Gender ratio (male:female)[a]	53:13		50:18		0.86	.35

Note. ADOS = Autism Diagnostic Observation Schedule; [a]Chi-square test

*$p < .001$.

Table 2

Means and Standard Deviations on CBCL/1.5-5 Scales for ASD and DD Groups

Scale	Group				
	ASD (*n* = 66) *M* (*SD*)	DD (*n* = 68) *M* (*SD*)	*t*	*p*	Effect size
Syndromes					
Emotionally reactive	61.09 (7.48)	59.00 (7.06)	1.66	.098	.287
Anxious/depressed	59.59 (9.93)	55.38 (5.60)	3.03	.003*	.522
Somatic complaint	56.67 (6.78)	57.16 (7.33)	-0.41	.686	-.069
Withdrawn	71.92 (11.24)	58.56 (7.91)	7.98	.000*	1.375
Sleep problems	56.53 (7.01)	55.94 (9.01)	0.42	.674	.0731
Attention problems	60.44 (7.48)	56.56 (6.59)	3.19	.002*	.550
Aggressive behavior	57.32 (7.25)	57.32 (8.98)	-0.00	.997	0
Broadband					
Internalizing	64.03 (8.93)	57.44 (9.35)	4.17	.000*	.721
Externalizing	57.61 (8.89)	55.28 (10.92)	1.35	.179	.234
Total problems	63.29 (9.70)	58.22 (11.00)	2.83	.005	.489
DSM-oriented					
Affective problems	61.29 (8.75)	58.12 (8.05)	2.18	.031	.377
Anxiety problems	60.79 (9.76)	58.76 (7.36)	1.36	.177	.235
PDP	72.68 (9.86)	63.51 (8.23)	5.85	.000*	1.010
Attention deficit/hyperactivity problems	58.99 (7.21)	58.43 (7.55)	0.44	.662	.076
Oppositional defiant problems	57.47 (8.14)	57.02 (8.23)	0.32	.748	.055

Note. PDP = Pervasive developmental problems; Effect sizes are Cohen's *d*.

* $p < .003$

Table 3

Number of Participants (%) Reaching the Borderline Range and Above for CBCL/1.5-5 Scales

Scale	Group		*p*	Effect size
	ASD (*n* = 66)	DD (*n* = 68)		
Syndromes[a]				
Emotionally reactive	26 (39.4%)	20 (29.4%)	.276	.105
Anxious/depressed	18 (27.3%)	7 (10.3%)	.015	.218
Somatic complaint	12 (18.2%)	16 (23.5%)	.526	-.066
Withdrawn	49 (74.2%)	16 (23.5%)	.000*	.507
Sleep problems	9 (13.6%)	7 (10.3%)	.602	.052
Attention problems	27 (40.9%)	10 (14.7%)	.001*	.293
Aggressive behavior	13 (19.7%)	16 (23.5%)	.677	-.047
Broadband[b]				
Internalizing	49 (74.2%)	34 (50.0%)	.005	.250
Externalizing	30 (45.5%)	23 (33.8%)	.216	-.119
Total problems	48 (72.7%)	32 (47.1%)	.003*	.262
DSM-oriented[a]				
Affective problems	21 (31.8%)	15 (22.1%)	.244	.110
Anxiety problems	18 (27.3%)	16 (23.5%)	.693	.043
PDP	55 (83.3%)	35 (51.5%)	.000*	.339

Attention deficit/hyperactivity problems	14 (21.2%)	18 (26.5%)	.546	-.062
Oppositional defiant problems	13 (19.7%)	15 (22.1%)	.835	-.029

Note. [a]T scores ≥ 65. [b]T scores ≥ 60.

* $p < .003$

Table 4

Structure Matrix Showing Loadings of CBCL/1.5-5 Scales on Discriminant Function

Scale	Function
Withdrawn	.93
PDP	.68
Internalizing	.49
Attention problems	.37
Anxious/depressed	.35

Table 5

Consistency Between Diagnosis and Withdrawn, PDP, and Internalizing at Different Cutoffs

Different Cutoff Items	AUCs	Sensitivity	Specificity	PPV	NPV
Withdrawn	.83				
60		57/66 (86.4%)	36/68 (52.9%)	57/89 (64.0%)	36/45 (80.0%)
62		55/66 (83.3%)	44/68 (64.7%)	55/79 (69.6%)	44/55 (80.0%)
66		49/66 (74.2%)	52/68 (76.5%)	49/65 (75.4%)	52/69 (75.4%)
70		41/66 (62.1%)	57/68 (83.8%)	41/52 (78.9%)	57/82 (69.5%)
PDP	.77				
67		53/66 (80.3%)	40/68 (58.8%)	53/81 (65.4%)	40/53 (75.5%)
70		49/66 (74.2%)	44/68 (64.7%)	49/73 (67.1%)	44/61 (72.1%)
71		43/66 (65.2%)	52/68 (76.5%)	43/59 (72.9%)	52/75 (69.3%)
Internalizing	.70				
60		49/66 (74.2%)	33/68 (50.0%)	49/83 (59.0%)	33/50 (66.0%)
61		46/66 (69.7%)	39/68 (57.4%)	46/75 (61.3%)	30/50 (60.0%)
62		42/66 (66.7%)	41/68 (63.2%)	42/67 (63.0%)	41/65 (63.1%)

Table 6

Means and Standard Deviations on items of CBCL/1.5-5 Withdrawn and PDP scales for ASD and DD Groups

Item of Scale	Group				
	ASD (*n* = 66)	DD (*n* = 68)			Effect
	M (*SD*)	*M* (*SD*)	*t*	*p*	size
Withdrawn & PDP					
4. Avoids looking others in the eye	1.30 (.70)	.49 (.56)	7.47	.000*	1.29
23. Doesn't answer when people talk to him/her	1.33 (.66)	.81 (.63)	4.70	.000*	0.81
67. Seems unresponsive to affection	.86 (.76)	.29 (.49)	5.13	.000*	0.89
70. Shows little affection toward people	.73 (.80)	.34 (.56)	3.26	.001*	0.57
98. Withdrawn, doesn't get involved with others	.85 (.79)	.15 (.40)	6.47	.000*	1.13
Withdrawn					
2. Acts young for age	.83 (.80)	.32 (.56)	4.28	.000*	0.74
62. Refuses to play active games	.56 (.59)	.27 (.51)	3.12	.002*	0.54
71. Shows little interest in things around him/her	.64 (.65)	.25 (.53)	3.78	.000*	0.65
PDP					
3. Afraid to try new things	.99 (.79)	.63 (.62)	2.87	.005	0.50
7. Can't stand things out of place	.61 (.68)	.63 (.71)	-.22	.827	-0.04
21. Disturbed by any change in routine	.59 (.68)	.47 (.59)	1.10	.273	0.19
25. Doesn't get along with other children	.82 (.72)	.41 (.60)	3.54	.001*	0.61
63. Repeatedly rocks head or body	.38 (.58)	.44 (.63)	-.60	.552	-0.10
76. Speech problem	1.55 (.73)	1.52 (.72)	.25	.806	0.04
80. Strange behavior	.59 (.78)	.18 (.52)	3.60	.000*	0.63
92. Upset by new people or situation	.88 (.75)	.66 (.77)	1.65	.101	0.29

Note. Effect sizes are Cohen's *d*

*$p < .003$

Table 7

Number of Participants (%) rated either 1 or 2 on items of CBCL/1.5-5 Withdrawn and PDP scales

Item of Scale	Group		*p*	Effect size
	ASD (*n* = 66)	DD (*n* = 68)		
	1&2	1&2		
Withdrawn & PDP				
4. Avoids looking others in the eye	57 (86.4%)	31 (45.6%)	.000*	.429
23. Doesn't answer when people talk to him/her	59 (89.4%)	47 (69.1%)	.004	.249
67. Seems unresponsive to affection	42 (63.6%)	19 (27.9%)	.000*	.358
70. Shows little affection toward people	34 (51.5%)	20 (29.4%)	.009	.225
98. Withdrawn, doesn't get involved with others	40 (60.6%)	9 (13.2%)	.000*	.492
Withdrawn				
2. Acts young for age	39 (59.1%)	19 (27.9%)	.000*	.314
62. Refuses to play active games	34 (51.5%)	16 (23.5%)	.001*	.289
71. Shows little interest in things around him/her	36 (54.5%)	14 (20.6%)	.000*	.351
PDP				
3. Afraid to try new things	45 (68.2%)	38 (55.9%)	.143	.127
7. Can't stand things out of place	33 (50.0%)	34 (50.0%)	1.000	.000
21. Disturbed by any change in routine	32 (48.5%)	29 (42.6%)	.498	.059
25. Doesn't get along with other children	42 (63.6%)	24 (35.3%)	.001*	.283
63. Repeatedly rocks head or body	22 (33.3%)	25 (36.8%)	.677	.036
76. Speech problem	57 (86.4%)	59 (86.8%)	.946	.006
80. Strange behavior	27 (40.9%)	8 (11.8%)	.000*	.332
92. Upset by new people or situation	43 (65.2%)	33 (48.5%)	.052	.168

Note. Effect sizes are Phi coefficient

*p<.003