

**Running head:** CHILD CARE WORKERS AS INFORMANTS IN SCREENING FOR  
ASD

**Can Child Care Workers Contribute to the Early Detection of Autism Spectrum  
Disorders? A Comparison between Screening Instruments with Child Care Workers  
versus Parents as Informants.**

Mieke Dereu, Ruth Raymaekers, Petra Warreyn, Inge Schietecatte, Mieke Meirsschaut, and

Herbert Roeyers

Ghent University

## **Abstract**

Several screening instruments for ASD in young children were developed during the last decades. Only few studies compare the discriminative power of these instruments in the same sample. In particular comparisons of instruments that use different informants are scarce in young children. The current study compared the discriminant ability of the Checklist for Early Signs of Developmental Disorders (CESDD) filled out by child care workers with that of frequently used parent questionnaires in a sample of 357 children between 5.57 and 48.13 months old who showed signs of ASD or language delay. The discriminant power of the CESDD was as good as that of parent questionnaires. Therefore, inclusion of child care workers in the early detection of ASD seems promising.

**Key words:** Autism Spectrum Disorder, Screening, Early identification, Child care, Toddlers

Mieke.Dereu@UGent.be

**Can Child Care Workers Contribute to the Early Detection of Autism Spectrum Disorders? A Comparison between Screening Instruments with Child Care Workers versus Parents as Informants.**

In recent years, the importance of early detection of autism spectrum disorders (ASD) has become more and more apparent. Early diagnosis of ASD enables early intervention which may improve developmental outcomes (Dawson, 2008; Rogers & Vismara, 2008). To enhance early detection, many screening instruments for ASD in young children were developed during the last two decades.

Some of these instruments use a combination of the clinical judgement of professionals such as community doctors and parental report. These professionals use the screening instrument to screen a population of children of a certain age to select those children who are at high risk for ASD. Examples of these screening instruments are the Checklist for Autism in Toddlers (CHAT; Baron-Cohen, Allen, & Gillberg, 1992; Baron-Cohen et al., 1996) and the 4-item Early Screening for Autistic Traits questionnaire pre-screen (ESAT; Dietz, Swinkels, van Daalen, van Engeland, & Buitelaar, 2006). These physicians have to base their clinical judgment on a brief observation of the child and a short conversation with the parents. This may possibly explain the low sensitivity reported for these screening instruments in the general population (Baird, et al., 2000; Groen, Swinkels, van der Gaag, & Buitelaar, 2007).

Other screening instruments for ASD in young children include parents as informants. Parents are seen as 'hands-on' experts who can best describe the typical behaviour of their child. Nevertheless, many parents have difficulties comparing their child to peers as they do not know which skills are expected at a certain age or may be biased in their answers and over- or under report problems in their child (Dumont-Mathieu & Fein, 2005). Examples of

these screening instruments are the 14-item ESAT (Dietz et al., 2006; Swinkels et al., 2006), the Modified Checklist for Autism in Toddlers (M-CHAT; Robins, Fein, Barton, & Green, 2001), the First Year Inventory (FYI; Reznick, Baranek, Reavis, Watson, & Crais, 2007) and the Social Communication Questionnaire (SCQ; Rutter, Bailey, & Lord, 2003). Most of the parent report measures that screen for ASD in young children revealed good psychometric properties (e.g., Robins et al., 2001; Watson et al., 2007), however most studies were based on referred children and not on a total population sample, except for the 14-item version of the ESAT (Dietz et al., 2006).

Recently, a new screening instrument that can be filled out by child care workers was developed to screen for ASD in the general population: the Checklist for Early Signs of Developmental Disorders (CESDD; Dereu et al., 2010). Child care workers have many opportunities to observe the children. They also have a good knowledge of the typical development of children and are able to compare across children with a similar developmental level (Branson, Vigil, & Bingham, 2008). The aim of the current study is to compare the discriminant ability of the CESDD with that of frequently used parent questionnaires such as the ESAT, M-CHAT and SCQ in a sample of toddlers.

Other studies have also compared the discriminant power of screening instruments for ASD. For example, Charman and colleagues (2007) compared the SCQ, the Children's Communication Checklist (CCC; Bishop, 1998) and the Social Responsiveness Scale (SRS; Constantino & Gruber, 2005) in a sample of 119 high-risk children. These children were however between 10 and 13 years old when the screening instruments were filled out. Only a few studies compared the discriminant ability of screening instruments at a very young age. Sikora, Hall, Hartley, Gerrard-Morris and Cagle (2008) compared scores on the Gilliam Autism Rating Scale (GARS; Gilliam, 1995) and the Child Behavior Checklist (CBCL; Achenbach & Rescorla, 2000) across ADOS-G classifications in 147 children between 36

and 71 months old. This study revealed low sensitivity and specificity for both GARS and the CBCL subscale 'withdrawn' (WD) in discriminating children with an autism classification on the ADOS-G from children with no classification. Only the CBCL subscale 'pervasive developmental problems' (PDP) had a satisfactory sensitivity of .80 but at the cost of a low specificity of .42. Overall, the GARS generated more false positive results than the CBCL-WD and CBCL-PDP subscales. A recent study by Oosterling and colleagues (2009) is so far the only study to offer an elaborated comparison of three screening instruments in very young children. They included 238 children between 8 and 44 months old at high risk for ASD derived from a large scale population screening with the ESAT. In their study, the ability of the ESAT, SCQ and the Communication and Symbolic Behavior Scales-Developmental Profile, Infant-Toddler Checklist (CSBS-DP; Wetherby & Prizant, 2002) to discriminate between ASD and non-ASD cases was compared and the instruments were also evaluated on item level. Receiver Operating Characteristic (ROC) analysis Area Under the Curve (AUC) values showed that the discriminant ability of all three instruments was unsatisfactory (AUC's were .58, .67 and .73 for respectively the ESAT, SCQ and CSBS-DP). In their high-risk sample, the authors found a good sensitivity for the ESAT (.88) and SCQ (.84), but at the cost of a very low specificity (.14 and .28 for respectively the ESAT and SCQ).

All of the abovementioned studies compared different versions of parent report measures: autism specific versus broader screening instruments. Studies that compared different informants used the same screening instrument, for example the Autism Spectrum Screening Questionnaire (ASSQ; Echlers, Gillberg, & Wing, 1999) completed by parents and teachers (Matilla et al., 2009). Other studies compared screening instruments for ASD with diagnostic instruments, for example the SCQ or the Children's Social Behavior Questionnaire (CSBQ; Hartman, Luteijn, Serra, & Minderaa, 2006) completed by parents versus Autism Diagnostic Observation Schedule (ADOS; Lord, Rutter, DiLavore, & Risi, 1999) and Autism

Diagnostic Interview-Revised (ADI-R; Rutter, Le Couteur, & Lord, 2003) administered by trained professionals (Corsello et al., 2007; de Bildt et al., 2009). The current study is the first to compare different autism specific screening instruments in a very young sample across different informants: parents versus child care workers.

## **Method**

### ***Participants***

The sample consisted of 357 children (219 boys and 138 girls) out of a larger screening study of 7,092 children for whom both the CESDD and parent questionnaires were filled out. Children were between 5.57 and 38.40 months old when the CESDD was filled out ( $M = 20.50$ ;  $SD = 5.78$ ) and between 8.20 and 48.13 months old when parent questionnaires were filled out ( $M = 23.23$ ;  $SD = 6.09$ ).

These children were considered to have an elevated risk for ASD or language problems based on their CESDD results: 206 children failed the language milestone most relevant for their age, 77 children screened positive for ASD on the CESDD (i.e. failed two or more red flags of ASD included in the checklist) and another 74 children failed both the last language milestone and screened positive for ASD. Children who screened positive for ASD on the CESDD were considered as a high-risk group and children who only failed the language items on the CESDD as a medium-risk group for ASD.

### ***Measures***

*Checklist for Early Signs of Developmental Disorders (CESDD; Dereu et al., 2010).* This instrument was developed to be easily filled out by child care workers in day-care centres for children between 3 and 39 months old. The objective of the instrument is to screen

children within the general population and to select children at risk for ASD who warrant further assessment. The instrument consists of 25 dichotomous items. Based on the age of the child, a different number of items needs to be considered: only the first 12 items for children younger than 12 months, the first 23 items for children between 12 and 24 months and all 25 items for children older than 24 months. If children fail two or more items, they screen positive for ASD. Preliminary results based on a population screening of 6,808 children revealed good psychometric properties: the CESDD had a sensitivity of .80 and a specificity of .94. The CESDD also incorporates items regarding the language development of the child (these items are however not included in the total score) and asks whether children were able to babble at 12 months, to speak single words at 16 months and to speak two-word phrases at 24 months or whether they have had any loss of language skills in the past.

*Early Screening of Autistic Traits (ESAT; Dietz et al., 2006; Swinkels et al., 2006).* The ESAT is a parent questionnaire and consists of 14 yes/no questions. Children who fail three or more items on the ESAT screen positive for ASD. This instrument has a low sensitivity of .23, but a very high specificity of .99 (Groen et al., 2007). These results are based on a very large population screen of over 30,000 children, whereas many other instruments report the psychometric properties of a high-risk sample of children, thereby overestimating the sensitivity.

*Modified Checklist for Autism in Toddlers (M-CHAT; Robins et al., 2001).* This 23-item parent report screener for ASD was developed for children of 24 months old, but can also be used with somewhat younger or older children as well (Pandey et al., 2008). Children screen positive for ASD on this questionnaire if they fail at least three items in total or two or more out of six critical items, derived from discriminant function analyses (Robins et al., 2001). A rescreen of 940 children out of the original sample at age 4, revealed a sensitivity

of .85 and a specificity of .93 (Dumont-Mathieu & Fein, 2005). In the current study, the Dutch translation of the M-CHAT was used (Dereu, Meirsschaut, Warreyn, & Roeyers, 2006).

*Social Communication Questionnaire (SCQ; Rutter, Bailey, & Lord, 2003)*. This is a 40-item parent report questionnaire. Children screen positive for ASD if they fail 15 or more items. There are two different forms of the SCQ: a ‘current’ version for children younger than 5 years and a ‘lifetime’ version for children older than 5, which incorporates both questions for the whole lifespan and questions about past behaviour when the child was 4 or 5 years old. This screening instrument was originally designed and validated for children above age 4. However, the ‘current’ version can also be used in children between 2 and 4 years. Lowering the cut-off score to 11 or more items in this age group has been suggested (Allen, Silove, Williams, & Hutchins, 2007; Corsello, et al., 2007). In this study, the Dutch translation of the SCQ ‘current’ was used (Warreyn, Raymaekers, & Roeyers, 2004) and the cut-off score was set on 11 or more items.

*Short-form versions of the MacArthur Communicative Development Inventories (CDIs Short-forms; Fenson et al., 2000)*. In the current study, a Dutch adaptation of the CDIs Short-forms were used: the N-CDIs Short-forms (Zink & Lejaegere, 2003). The N-CDIs Short-forms were revalidated for Dutch speaking children in Belgium between 8 and 30 months. The N-CDI 1 was developed for children between 8 and 16 months old and the N-CDI 2A or 2B for children between 16 and 30 months old. In this study, only the versions N-CDI 1 or N-CDI 2A were used to measure the receptive and expressive language level of a child. Children who score below the 10<sup>th</sup> percentile and show at least 4 months delay on both receptive and expressive language, warrant further follow-up (Zink & Lejaegere, 2003).

[INSERT FIGURE 1 HERE]

### ***Procedure***

As part of a large scale screening study (see Figure 1, see also Dereu et al., 2010) in Flemish day-care facilities, 7,092 children between 3.00 and 38.90 months ( $M_{age} = 16.81$ ,  $SD = 8.17$ ) were screened for ASD with the CESDD by child care workers. The parents of 1,102 children ( $M_{age} = 20.64$ ,  $SD = 6.45$ , range 3.60 – 38.90) were asked to fill out additional parent questionnaires: the ESAT for all children accompanied by the FYI for children between 11 and 13 months, the M-CHAT for children between 18 and 24 months and the SCQ for children older than 24 months. In addition, parents were asked to fill out the N-CDIs Short-forms. The parents of 357 children filled out these questionnaires when their child was about 23.23 months old ( $SD = 6.09$ , range 8.20 – 48.13), on average 2.73 months ( $SD = 2.83$ ) after the CESDD was filled out for their child (see also Table 1 for demographics of the different screening instruments). Some of these parents already took part in another study on the Dutch translation of the M-CHAT (unpublished) ( $n = 57$ ). For these children, the parents were always given the SCQ as a second screener for ASD next to the ESAT, resulting in 27 children for whom the SCQ was filled out before they were 24 months old. Because we obtained only 29 FYI's for the youngest children, data on the FYI were omitted for the current study.

[INSERT TABLE 1 HERE]

As shown in Figure 1, the parents of all children with a positive screen and/or suspected language delay on the CESDD received the parent questionnaires. So, the subjects of the current study are all at risk for ASD or language problems. Day-care centres were asked to fill out the CESDD again after at least 3 months for children with a positive screen for ASD. Non compliant parents of screen positive children were repeatedly asked to fill out additional questionnaires if the child still showed signs of ASD or failed the language items at a subsequent screen with the CESDD. The return rates for the parent questionnaires are shown in Figure 2 and represent the compliance with the overall design, because parents who

already filled out questionnaires, received on subsequent positive screens no more questionnaires. Also, because the preliminary cut-off for positive screen on the CESDD was set on 4 instead of 2 signs for children above 12 months, the parents of 122 children older than 12 months who never failed the language items and who showed 2 or 3 signs of ASD on one or more occasions, were never given additional questionnaires. The return rates were corrected accordingly and represent the true return rates for parents that actually received questionnaires.

Children for whom the parent questionnaires were filled out and children for whom the parents were non-compliant did not differ significantly in gender, age of the child when parents were first given questionnaires and multilingual upbringing (all  $p > .05$ ).

[INSERT FIGURE 2 HERE]

A subsample of children ( $n = 70$  or 19.67%) was also seen for further assessment at the university lab with the ADOS (Lord et al., 1999) and the Mullen Scales of Early Learning (MSEL; Mullen, 1995). For demographics and scores on the ADOS-G and MSEL of children assessed by the research team, see Table 2.

[INSERT TABLE 2 HERE]

Children received further assessment when they were about 28.75 months old ( $SD = 7.02$ , range 13.47 – 51.40), on average 3.20 months ( $SD = 4.68$ ) after the parent questionnaires were filled out. As shown in Figure 1, all children with a positive screen on one or more screeners for ASD and/or language delay on the N-CDI, were invited for further evaluation. In addition, children who screened negative for ASD and had no suspected language delay on the parent questionnaires, but screened repeatedly positive on the CESDD were invited for the developmental assessment. Moreover, some children did not warrant further evaluation but were however seen at the university lab on request of their worried

parents. Figure 3 illustrates the parental compliance with the invitation for further evaluation of their child.

Children who were seen for evaluation at the university lab did not differ significantly from children for whom the parents declined further assessment in gender, age of the child when parents were first given questionnaires, multilingual upbringing, positive screens on the M-CHAT or SCQ and language delay on the N-CDI (all  $p > .05$ ). However, the children who were seen for further evaluation were somewhat older when parent questionnaires were completed compared to children who were not seen at the university lab (mean age 25.36 and 22.69 months respectively;  $t(130) = -2.205$ ,  $p = .029$ ). Also, the proportion of children with a positive screen for ASD on the CESDD and on the ESAT was higher in the compliant group than in drop-out children (80.95% and 62.32% respectively on the CESDD with  $\chi^2(1) = 5.578$ ,  $p = .021$ ; 27.12% and 8.70% respectively on the ESAT with  $\chi^2(1) = 7.584$ ,  $p = .009$ ).

[INSERT FIGURE 3 HERE]

Finally, based on the results of the ADOS-G, MSEL and the clinical judgment of the research team, children were referred for a diagnostic evaluation if necessary.

Diagnosis of ASD or other developmental disorders was made by independent university based autism clinics or diagnostic centres for developmental disorders across Flanders. In some children, this diagnosis was already made prior to the screening procedure ( $n = 12$ ; these children were not seen at the university lab but the day-care centres and parents informed us of the diagnosis), other children received a diagnosis after referral or were considered as developmentally or language delayed based on their MSEL results (scores below the 10<sup>th</sup> percentile) or suspected of ASD according to the ADOS algorithm and the clinical judgment of the research team ( $n = 56$ ).

Looking to the different diagnoses into more detail, in the total sample of 357 children 26 children were diagnosed with ASD and 7 children were highly suspected of ASD (sASD):

3 children received a working diagnosis in an autism clinic and currently await re-evaluation at an older age and another 4 children were suspected of ASD based on ADOS, MSEL and clinical judgment of the research group, but parents declined referral to an autism clinic or are still on a waiting list. Another 35 children received a diagnosis of other developmental disorders or had a developmental delay (DD): 3 children were diagnosed with mental retardation, 1 child with severe motor dysfunction, 1 child with a language disorder and 11 children with general developmental delay; also 5 children were considered as developmentally delayed based on their MSEL scores (i.e. Early Learning Composite scores below the 10<sup>th</sup> percentile) and 14 children showed language delay on the MSEL (i.e. scores on expressive and/or receptive language below the 10<sup>th</sup> percentile).

### *Data analysis*

For each screening instrument, total scores were compared across diagnostic groups using the Kruskal-Wallis Test. Pairwise comparisons between diagnostic groups and between high- and medium-risk groups based on the CESDD (i.e. children who screened positive for ASD or who only showed language delay on the CESDD, respectively) were obtained with the Mann-Whitney U Test. Non-parametric tests were chosen because of the skew distribution of the screening instruments used in this sample.

To compare the discriminant ability of the different screening instruments in distinguishing children with ASD from those without ASD, different measures were calculated: Sensitivity (Se), Specificity (Sp), Positive Predictive Value (PPV) and Negative Predictive Value (NPV). For comparing Se and Sp between instruments, the McNemar  $\chi^2$  test was used (Trajman, & Luiz, 2008). The method of generalized estimating equations (GEE) was used to fit a logit link function to the data and PPVs and NPVs were then compared using a Wald test (Leisenring, Alonzo, & Pepe, 2000). Also ROC analyses were performed to

assess and compare the discriminant power of the different instruments using the StAR server (Vergara, Norambuena, Ferrada, Slater, & Melo, 2008). The discriminant ability of all individual items in the different screening instruments in distinguishing ASD from non-ASD cases was assessed with Chi Square analyses and Phi-coefficients.

## Results

### *Comparisons of Mean Total Scores*

For all parent questionnaires, screening for ASD, high-risk children (children who screened positive for ASD on the CESDD) had a significantly higher total score than medium-risk children (children who only failed the language items on the CESDD) (see Table 3). Also, the Kruskal-Wallis Test showed that for all screening instruments the mean total scores differed significantly across diagnostic groups ( $\chi^2(3)$  ranged from 29.33 to 70.50, all  $p < .001$ ). Pairwise comparisons of the diagnostic groups for the CESDD, ESAT, M-CHAT and SCQ are shown in Table 4.

[INSERT TABLE 3 AND 4 HERE]

On all screening instruments, children with ASD scored significantly higher than children who did not receive a diagnosis. They also scored significantly higher than children with developmental delay, except on the M-CHAT. On all questionnaires, there was no significant difference between children highly suspected of having ASD and children with ASD. In further analyses, these diagnostic groups were combined into ASD cases. Also, children highly suspected of ASD scored significantly higher than children with no diagnosis and with developmental delay, except on the SCQ. In addition, on all instruments the scores of children with developmental delay were significantly higher than the scores of children with no diagnosis.

### *Discriminant Ability of the Screening Instruments*

To compare the discriminant power of the different screening instruments, two groups of children were formed: one group of children for whom the CESDD, ESAT and M-CHAT were filled out ( $n = 197$ ) and one group of children for whom the CESDD, ESAT and SCQ were filled out ( $n = 149$ ). There was some overlap between these groups because for 57 children who also participated in a validation study of the Dutch translation of the M-CHAT (unpublished), both M-CHAT and SCQ were available. Children for whom only the ESAT was filled out or whose data were incomplete ( $n = 68$ ), were omitted from these analyses. The indices of diagnostic accuracy in distinguishing between ASD cases (i.e. children highly suspected of ASD or with a clinical diagnosis of ASD) and non-ASD cases (i.e. children with other developmental difficulties or no diagnosis) for the different screening instruments are represented in Table 5 and Table 6.

[INSERT TABLE 5 AND 6 HERE]

ROC analyses showed that the discriminant power of the CESDD is at least as good as that of parent questionnaires. The AUCs of the different screening instruments did not differ significantly from each other ( $p$ -values ranged from .068 to .853 in pairwise comparison). However, in the group of children for which the CESDD, ESAT and SCQ were filled out there was a trend towards a higher AUC for the CESDD in comparison to the ESAT ( $p = .068$ ). For all questionnaires used in this study, the AUCs fell between .70 and .90, indicating a moderate to good concordance between the scores on the screening instruments and the clinical diagnosis (Akobeng, 2007). Also, the CESDD had the highest Se, but this came at the cost of the lowest Sp. Moreover, the Se of the CESDD was significantly higher than that of the ESAT ( $p = .016$  and  $p = .001$ ), but the Sp of the CESDD was significantly lower than that of all parent questionnaires used in both groups of children (all  $p \leq .001$ ).

The PPV and NPV of the CESDD were comparable to that of the M-CHAT and the SCQ (all  $p > .05$ ). Only the ESAT had a significantly higher PPV than the CESDD ( $p = .018$  and  $p = .005$ ) but at the cost of a much lower Se. In addition, the NPV of the CESDD was (marginal) significantly higher than that of the ESAT ( $p = .055$  and  $p = .021$ ).

### ***Item Analyses***

To assess the discriminant power of every item in the parent questionnaires and the CESDD, Chi-square analyses were performed and Phi-coefficients were calculated (see Table 7). Phi can have a value ranging from -1.00 to +1.00 and can be interpreted like Pearson correlation coefficients with values between .10 and .30 indicating a weak association, values between .30 and .60 indicating a moderate association and values above .60 indicating strong association (Agresti, 2007).

For all questionnaires, items regarding *joint attention skills* are amongst the items with the highest discriminant power based on the Phi-coefficient. Items regarding ‘lack of or absence of showing behaviour’ had Phi-values of .34, .32 and .37 on respectively the CESDD, ESAT and M-CHAT. The item regarding ‘response to joint attention’ was amongst the items with the highest Phi-coefficient on the CESDD (.36). In addition, ‘no pointing’ was amongst the items with the highest phi-coefficient on the M-CHAT (.44 for imperative and .40 for declarative pointing) and SCQ (.35).

Many items regarding *play development* had Phi-values of .30 or above: ‘lack of functional play’ (.30), ‘lack of symbolic play’ (.38) or ‘stereotyped repetitive play’ (.39) on the CESDD, ‘no varied play’ on the ESAT (.33), and ‘no imitative social play’ (.39), ‘no imaginative play with peers’ (.38) and ‘no group play’ (.40) on the SCQ. For the M-CHAT, none of the items regarding play development had Phi-values of .30 or above. The content of

these items are however very different according to the age range for which the instruments were developed.

In addition to items regarding joint attention and play development, some items regarding *social orientation* had high Phi-coefficients: ‘lack of social smile’ (.32) and ‘use of someone’s hand as an instrument’ (.40) on the CESDD, ‘not easy to make eye contact’ (.35) and ‘no interest in children or adults’ (.35) on the ESAT, and also ‘no interest in children’ (.42) and ‘no attention to voice’ (.40) on the SCQ.

Items regarding *sensory abnormalities* had also high Phi-values. The item on the SCQ asking parents about ‘unusual sensory interests’ of their child even had a Phi-value of .50. The ESAT asks parents about ‘unusual reaction to sensory stimuli’ and had a Phi-value of .31.

Finally, the item on the CESDD assessing ‘lack of *imitation*’ also had a Phi-value of .30.

### ***Evaluation of the screening design***

Out of the 33 children with ASD or suspected ASD, 28 children were invited for further assessment at the university lab because of developmental concerns based on the different questionnaires used (see Figure 1; for a complete description of the screening procedure used: see Dereu et al., 2010). The parents of 4 children with ASD declined this invitation to further participate in the study because their child was already diagnosed elsewhere. The five children with ASD who did not warrant further assessment based on the results of the parent questionnaires, were however all seen for further assessment on request of the parents because the child had a sibling with ASD ( $n = 3$ ) or because another professional had formulated concerns about the development of the child ( $n = 2$ ).

Only 4 children with (suspected) ASD screened false negative on the CESDD, whereas 9 children screened negative on all available parent report screening instruments for

ASD (i.e. ESAT and M-CHAT and/or SCQ). Only two children with (suspected) ASD screened negative on all screening instruments filled out for that child (so both on child care worker as parent report). So, the combination of the CESDD with different parent questionnaires for screening for ASD detected 93.94% of children with (suspected) ASD. However, 177 children screened positive on at least one of the combined screening instruments (CESDD and parent questionnaires), resulting in 82.49% false positives for ASD and an overall accuracy of 58.54% correctly classified cases.

## **Discussion**

The aim of the current study was to compare the discriminative validity of a screening instrument based on child care workers' observations with that of well known parent questionnaires to screen for ASD in very young children.

ROC analyses showed that all the screening instruments in the current study had a satisfactory discriminant power based on the AUCs. Furthermore, the CESDD could discriminate as well as parent questionnaires between children with and without ASD. However, when we compare the indices of diagnostic accuracy of the different screening instruments, none of the instruments had satisfactory values for both Se and Sp (Meisels (1989) stated that both Se and Sp should at least be .80). The CESDD was the only instrument with an excellent Se, but at the cost of a fair or even poor Sp, depending on the age group for which the CESDD was filled out (Cicchetti, Volkmar, Klin, and Showalter (1995) established the clinical significance of indices of diagnostic accuracy as: <.70 = poor; .70-.79 = fair; .80-.89 = good; .90-1.00 = excellent). Moreover, the Se of the CESDD was significantly better than that of the ESAT, but the Sp of the CESDD was lower than that of all parent questionnaires used.

All instruments had a good NPV but a poor PPV, reflecting the many false positive results in the current sample, especially in the younger age group (for which the CESDD, ESAT and M-CHAT were filled out). Some of these false positive results may be found in children with other developmental problems. The PPV and NPV of the CESDD were comparable with those of the M-CHAT and SCQ. The ESAT had a significantly higher PPV in both age groups, but also a significantly lower NPV in the older age group (for which the CESDD, ESAT and SCQ were filled out) compared to the CESDD .

No instrument seemed to perform better than the other instruments in general. Therefore the choice for a screening instruments will depend on the purpose of the screening. Glover and Albers (2007) stated that when the consequences associated with underidentification are great (i.e. children may not receive much-needed supports and services), it may be useful to compromise precision (e.g., a high PPV or Sp) for inclusion (Se). Especially in screening protocols with subsequent stages, sensitivity in the first stage is crucial for insuring that no children who are potentially at risk are overlooked. So, in a first stage, one may opt to use the CESDD because of its excellent Se. Increasing positive predictive value is expected for subsequent stages. Then the ESAT seems a good option based on its higher Sp and PPV. For the current sample, this particular combination would have resulted in 92.63% of the children correctly classified and an increased PPV of .63. However, the Se would still have been poor (.39). Based on the current study, it is not possible to present the best combination for overall screening. Depending on setting and purpose of screening, other combinations will be more preferential.

One should interpret these findings with caution for several reasons. First of all, the sample largely consisted of children who screened positive on the CESDD (the high-risk group). The other children all had language delay. So, this sample consisted of developmentally challenged children instead of a general population sample. The current

study design (inclusion based on positive screen and/or language delay on the CESDD) may have resulted in an overestimation of the Se but also in an underestimation of the Sp of the CESDD. This may also partly explain differences found in Se, Sp, PPV and NPV for the parent questionnaires in comparison to other studies. For example, Oosterling and colleagues (2009) also compared the discriminant ability of the ESAT and SCQ in 238 children between 8 and 44 months old. However, their sample was mainly selected based on a positive screen on the ESAT (only 30 children were included based on clinical concern). This resulted in a much higher Se of .88 for the ESAT, compared to the calculated Se of .23 by Groen and colleagues (2007) that was based on the prevalence rate of ASD and the results of a large population screen with the ESAT. In our study, the Se of the ESAT was estimated at .38. Also, the study of Oosterling and colleagues (2009) had higher estimates for the PPV and much lower estimates for the Sp, NPV and AUC compared to our study, for both the ESAT and the SCQ. Again, this can be explained by the inclusion criteria in the different studies: based on positive screen on the parent questionnaire ESAT versus a positive screen or language delay on the CESDD filled out by a child care worker. A second reason why caution in interpreting these results is warranted, is the rather low amount of ASD cases. This may have biased the estimates of Se, Sp, PPV and NPV. Furthermore, when statistically comparing Se, the McNemar  $\chi^2$  test is only based on children with ASD or suspected ASD. So these tests were performed on very small sample sizes.

We assessed the discriminant ability of the items in the screening instruments but we chose not to report the Se, Sp, PPV and NPV for each item because of the low number of children with ASD or suspected ASD in our sample. However, we reported the percentage of children with and without ASD for which each sign of ASD was recognized.

Looking at the content of the items with the highest discriminant power based on the Phi – coefficient, this were mainly items regarding joint attention skills. However, there are

great differences in the formulation of the items and consequently in the percentage of children with ASD for which the item was recognized. For example, the CESDD incorporates an item about 'lack of showing objects to others to indicate interest'. This item was checked for 64.52% of children with ASD, whereas the item regarding 'absence of showing' in the ESAT and M-CHAT was recognized in only 26.67% and 28.57% of children, respectively. A similar item regarding 'no showing' in the SCQ was only recognized in 13.04% (reflected by a much lower Phi-value of .15), possibly due to the phrasing of the item. If a child has ever shown something, the child passes this item, even if it occurred only once in the lifespan of the child. This child would fail the item in the CESDD, because of a quantitative difference between this child and its peers who regularly show objects. Another possible explanation is the higher age of the children for which the SCQ was filled out: it is possible that some children learned how to show objects by the time the SCQ was filled out.

Also, one should take into account the difference between the informants used: child care workers know better what should be expected of children at a certain developmental level and can possibly easier recognize subtle qualitative or quantitative impairments in these social-communicative behaviours. Moreover, prior to participation in the screening study, child care workers were trained in recognizing these early signs of ASD, possibly resulting in the fact that the CESDD is the only instrument for which all but one item significantly discriminated between children with and without ASD at the alpha .01 significance level.

Next to showing behaviours, items regarding 'pointing to indicate interest' were among the items that best discriminated between children with and without ASD. Only the ESAT does not incorporate an item regarding 'declarative pointing'. On the CESDD, the item regarding 'lack of pointing' was recognized in 48.39% of children with ASD and had a Phi-value of .28. Similar items in the M-CHAT and SCQ regarding 'absence of pointing

behaviours to express interest' were recognized in respectively 42.86% and 34.78% of children with ASD. Phi-values were .40 on the M-CHAT and .35 on the SCQ.

In addition, items regarding 'response to joint attention' on the CESDD and the M-CHAT (such items are not found in the ESAT and SCQ) had high Phi – coefficients. Other items with high Phi – values included items regarding play development, imitation, social orienting and sensory abnormalities.

### ***Limitations***

Although the sample was comprised of a risk group, the inclusion of a medium-risk group resulted in a relatively low proportion of children with (suspected) ASD in the sample. Especially in the group children for which the CESDD, ESAT and M-CHAT were filled out, the amount of children with ASD or suspected ASD was quite low. The estimation of the psychometric properties of the different screening instruments should be interpreted with caution. In particular the comparison of the Se between screening instruments was based on small sample sizes. Moreover, when total scores for different diagnostic groups were compared, children with ASD and children suspected of ASD were analyzed separately and mainly the amount of children suspected of ASD was rather low.

Also, because the sample constituted of children who screened positive for ASD or failed the language items on the CESDD, this could have biased the results in favour of the CESDD in comparison to the other screening instruments. Replication of the findings in a more selected sample with a higher proportion of children with ASD is necessary. Ideally would be replication in a very large population study with a sufficient proportion of children with ASD. Only then the estimates of Se, Sp, PPV and NPV of these questionnaires can be compared reliably for use in the general population .

Furthermore, the findings for the ESAT, M-CHAT and SCQ should be interpreted with some caution. In this study, the ESAT was filled out by the parents alone. In the original ESAT population screening, the ESAT was filled out by a child psychologist during a 1.5-hour home visit. When either the parent or the child psychologist found the behaviour of the child on an item atypical, the child screened positive for ASD on that item (Dietz et al., 2006). In addition, positive screens on the M-CHAT are based solely on parent report. We did not include the telephone interview for positive screen results, as suggested to lower the false positive rate (Dumont-Mathieu & Fein, 2005). Finally, the SCQ was used in very young children, whereas this instrument was originally developed for children aged 4 or above (Rutter et al., 2003). This could possibly explain the lower Phi-values for these items and suggests that one should rather rely on a total score than on separate items.

Incorporation of children with language delay in further phases of the screening protocol allowed to detect some missed cases on the CESDD and to evaluate how these children performed on the other screening instruments. However, we have no insight in possible missed cases with no language delay. We are currently conducting a survey in all university based autism clinics in Flanders and ask all parents of children diagnosed with ASD in the past few years in Flanders to inform us if their child was ever in a day-care centre that participated in the screening study. In addition, we are sending questionnaires to a subsample of day-care centres and ask the parents of the children who participated in the screening study retrospectively to inform us of the diagnostic status of their child. So far, no diagnosis of ASD was missed in the current sample of 357 children, but the survey is still ongoing and will have to be repeated in the next years to obtain final results.

Finally, return rates for parent questionnaires were rather low and only about half of the children invited for further assessment were seen at the university lab (the parents of the other children declined our invitation). Children of compliant and non-compliant parents

were compared in order to assess the generalizability of the results of this study. They did not differ from each other in gender, age at which additional parent questionnaires were first given and multilingual upbringing. However, the children that were seen for further assessment at the university lab were somewhat older when parents filled out the additional questionnaires and the proportion of positive screens on the CESDD and on the ESAT were higher in the compliant group compared to children of parents that declined further assessment for their child. A closer look at the response rates for both the parent questionnaires and further assessment showed that they were highest in children with both language delay and signs of ASD and lowest in children with language delay only. It seems that parents were more inclined to participate in further stages of the screening design when the atypical development of their child was more apparent. Similarly, Dietz, Swinkels, van Daalen, van Engeland, and Buitelaar (2007) found for their ESAT-screening study that the parental compliance was higher in children who were already identified as at risk for ASD by physicians in well-baby clinics in comparison to the compliance in a general population screening for ASD. However, it remains possible that many parents are not yet willing to respond to concerns about their child's development raised by others such as physicians or child care workers, resulting in low response rates.

### ***Conclusion***

This study reports on the psychometric properties of screening instruments used in a screening protocol that incorporates parents as well as child care workers as informants. Results of ROC analyses showed that the CESDD based on the report of child care workers performed equally well as parent report screening instruments in the early detection of ASD.

Due to the screening protocol with very low thresholds, all children with ASD or suspected ASD were seen or at least invited for further assessment. However, only 41.43% of

all the children that were seen at the university lab got an ASD (working) diagnosis after referral. Another 27 children warranted further follow up because of language delay, delayed motor development or general developmental delay based on the MSEL.

Incorporating the report of child care workers in the screening for ASD seems very useful to optimize the early detection of ASD. Moreover, only 4 out of the 33 children with (suspected) ASD falsely screened negative on the CESDD, whereas 9 children screened negative on all available parent report screening instruments for ASD (i.e. ESAT and M-CHAT and/or SCQ). Combining different screening instruments with different informants can help in detecting more children with possible ASD: 93.94% of the children screened positive on at least one screening instrument used. However, when all children who screened positive on either one of the three screening instruments would have been considered as at risk for ASD, this would have resulted in 82.49% false positive cases.

One should rather consider the CESDD as a Level one population screener within day-care facilities. The instrument can select children with possible developmental problems for whom parent screeners for ASD can be filled out next. That way only the parents of children at risk for developmental problems who warrant further assessment are involved in further stages, without alarming the parents of the typically developing children. What combination of instruments is best used, cannot be determined based on the results of this study. Furthermore, the setting of the screening, the age of the children and the preferred threshold for further evaluation, will influence this choice and will make different screening instruments and combinations more preferential.

## References

- Achenbach, T., & Rescorla, L. (2000). *Child Behavior Checklist*. Burlington, Vermont: Achenbach System of Empirically Based Assessment.
- Agresti, A. (2007). *An introduction to categorical data analysis*. New York: Wiley.
- Akobeng, A. K. (2007). Understanding diagnostic tests 3: receiver operating characteristic curves. *Acta Paediatrica*, *96*, 644-647. doi: 10.1111/j.1651-2227.2006.00178.x.
- Allen, C. W., Silove, N., Williams, K., & Hutchins, P. (2007). Validity of the Social Communication Questionnaire in assessing risk of autism in preschool children with developmental problems. *Journal of Autism and Developmental Disorders*, *37*, 1272-1278. doi: 10.1007/s10803-006-0279-7.
- Baird, G., Charman, T., Baron-Cohen, S., Cox, A., Swettenham, J., Wheelwright, S., & Drew, A. (2000). A screening instrument for autism at 18 months of age: A 6-year follow-up study. *Journal of the American Academy of Child & Adolescent Psychiatry*, *39*, 694-702.
- Baron-Cohen, S., Allen, J., & Gillberg, C. (1992). Can autism be detected at 18 months? The needle, the haystack, and the CHAT. *British Journal of Psychiatry*, *161*, 839-843.
- Baron-Cohen, S., Cox, A., Baird, A., Swettenham, J., Nightingale, N., Morgan, K., Drew, A., & Charman, T. (1996). Psychological markers in the detection of autism in infancy in a large population. *British Journal of Psychiatry*, *168*, 158-163.
- Bishop, D. V. M. (1998). Development of the Children's Communication Checklist (CCC): A method for assessing qualitative aspects of communicative impairment in children. *Journal of Child Psychology and Psychiatry*, *39*, 879-891.

- Branson, D., Vigil, D. C., & Bingham, A. (2008). Community childcare providers' role in the early detection of autism spectrum disorders. *Early Childhood Education Journal*, *35*, 523-530. doi: 10.1007/s10643-008-0243-6.
- Charman, T., Baird, G., Simonoff, E., Loucas, T., Chandler, S., Meldrum, D., & Pickles, A. (2007). Efficacy of three screening instruments in the identification of autistic-spectrum disorders. *British Journal of Psychiatry*, *191*, 554-559. doi: 10.1192/bjp.bp.107.040196.
- 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.
- Constantino, J. N., & Gruber, C. P. (2005). *Social Responsiveness Scale (SRS)*. Los Angeles: Western Psychological Services.
- Corsello, C., Hus, V., Pickles, A., Risi, S., Cook, E. H., Leventhal, B. L., & Lord, C. (2007). Between a ROC and a hard place: decision making and making decisions about using the SCQ. *Journal of Child Psychology and Psychiatry*, *48*, 932-940. doi: 10.1111/j.1469-7610.2007.01762.x.
- Dawson, G. (2008). Early behavioral intervention, brain plasticity, and the prevention of autism spectrum disorder. *Development and Psychopathology*, *20*, 775-803. doi: 10.1017/S0954579408000370.
- de Bildt, A., Mulder, E. J., Hoekstra, P. J., van Lang, N. D.J., Minderaa, R. B., & Hartman, C. A. (2009). Validity of the Children's Social Behavior Questionnaire (CSBQ) in children with intellectual disability: Comparing the CSBQ with the ADI-R, ADOS, and clinical DSM-IV-TR classification. *Journal of Autism and Developmental Disorders*, *39*, 1464-1470. doi: 10.1007/s10803-009-0764-x.

- Dereu, M., Meirsschaut, M., Warreyn, P., & Roeyers, H. (2006). *The Modified Checklist for Autism in Toddlers. Dutch translation*. Retrieved March 30, 2010, from Georgia State University, Department of Psychology Website: [http://www2.gsu.edu/~psydlr/Diana\\_L.\\_Robins,\\_Ph.D.\\_files/M-CHAT\\_Dutch.pdf](http://www2.gsu.edu/~psydlr/Diana_L._Robins,_Ph.D._files/M-CHAT_Dutch.pdf).
- Dereu, M., Warreyn, P., Raymaekers, R., Meirsschaut, M., Pattyn, G., Schietecatte, I., & Roeyers, H. (2010). Screening for autism spectrum disorders in Flemish day-care centres with the Checklist for Early Signs of Developmental Disorders. *Journal of Autism and Developmental Disorders*, *40*, 1247-1258. doi: 10.1007/s10803-010-0984-0.
- Dietz, C., Swinkels, S., van Daalen, E., van Engeland, H., & Buitelaar, J. K. (2006). Screening for autistic spectrum disorder in children aged 14-15 Months. II: Population screening with the Early Screening of Autistic Traits questionnaire (ESAT). Design and general findings. *Journal of Autism and Developmental Disorders*, *36*, 713-722. doi: 10.1007/s10803-006-0114-1.
- Dietz, C., Swinkels, S. H. N., van Daalen, E., van Engeland, H., & Buitelaar, J. K. (2007). Parental compliance after screening social development in toddlers. *Archives of Pediatrics and Adolescent Medicine*, *161*, 363-368.
- Dumont-Mathieu, T., & Fein, D. (2005). Screening for autism in young children: The Modified Checklist for Autism in Toddlers (M-CHAT) and other measures. *Mental Retardation and Developmental Disabilities*, *11*, 253-262. doi: 10.1002/mrdd.20072.
- Ehlers, S., Gillberg, C., & Wing, L. (1999). A screening questionnaire for Asperger syndrome and other high-functioning autism spectrum disorders in school age children. *Journal of Autism and Developmental Disorders*, *29*, 129-141.

- Fenson, L., Pethick, S., Renda, C., Cox, J. L., Dale, P. S., & Reznick, J. S. (2000). Short-form versions of the MacArthur Communicative Development Inventories. *Applied Psycholinguistics, 21*, 95-116.
- Gilliam, J. E. (1995). *Gilliam Autism Rating Scale*. Texas: Pro-ed.
- Glover, T. A., & Albers, C. A. (2007). Considerations for evaluating universal screening assessments. *Journal of School Psychology, 45*, 117-135. doi: 10.1016/j.jsp.2006.05.005.
- Groen, W. B., Swinkels, S. H., van der Gaag, R. J., & Buitelaar, J. K. (2007). Finding effective screening instruments for autism using Bayes theorem. *Archives of Pediatrics and Adolescent Medicine, 161*, 415-416.
- Hartman, C. A., Luteijn, E., Serra, M., & Minderaa, R. (2006). Refinement of the Children's Social Behavior Questionnaire (CSBQ): An instrument that describes the diverse problems seen in milder forms of PDD. *Journal of Autism and Developmental Disorders, 36*, 325-342. doi: 10.1007/s10803-005-0072-z.
- Hollingshead, A.B. (1975). *Four factor index of social status*. Unpublished manuscript, Yale University.
- Leisenring, W., Alonzo, T., & Pepe, M. S. (2000). Comparisons of predictive values of binary medical diagnostic tests for paired designs. *Biometrics, 56*, 345-351.
- Lord, C., Rutter, M., DiLavore, P., & Risi, S. (1999). *Autism Diagnostic Observation Schedule: Manual*. Los Angeles: Western Psychological Services.
- Matilla, M., Jusilla, K., Kuusikko, S., Kielinen, M., Linna, S., Ebeling, H., Bloigu, R., Joskitt, L., Pauls, D., & Moilanen, I. (2009). When does the Autism Spectrum Screening Questionnaire (ASSQ) predict autism spectrum disorders in primary school-aged children? *European Child and Adolescent Psychiatry, 18*, 499-509. doi: 10.1007/s00787-009-0044-5.

- Meisels, S. J. (1989). Can developmental screening tests identify children who are developmentally at risk? *Pediatrics*, *83*, 578-585.
- Mullen, E. M. (1995). *Mullen Scales of Early Learning*. Circle Pines MC: American Guidance Service.
- Oosterling, I. J., Swinkels, S. H., van der Gaag, R. J., Visser, J. C., Dietz, C., & Buitelaar, J. K. (2009). Comparative analysis of three screening instruments for autism spectrum disorder in toddlers at high risk. *Journal of Autism and Developmental Disorders*, *39*, 897-909. doi: 10.1007/s10803-009-0692-9.
- Pandey, J., Verbalis, A., Robins, D. L., Boorstein, H., Klin, A., Babitz, T., Chawarska, K., Volkmar, F., Green, J., Barton, M., & Fein, D. (2008). Screening for autism in older and younger toddlers with the Modified Checklist for Autism in Toddlers. *Autism*, *12*, 513-535. doi: 10.1177/1362361308094503.
- Reznick, J. S., Baranek, G. T., Reavis, S., Watson, L. R., & Crais, E. R. (2007). A parent-report instrument for identifying one-year-olds at risk for an eventual diagnosis of autism: The First Year Inventory. *Journal of Autism and Developmental Disorders*, *37*, 1691-1710. doi: 10.1007/s10803-006-0303-y.
- Robins, D. L., Fein, D., Barton, M. L., & Green, J. A. (2001). The Modified Checklist for Autism in Toddlers: An initial study investigating the early detection of autism and pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, *31*, 131-144.
- Rogers, S. J., & Vismara, L. A. (2008). Evidence-based comprehensive treatments for early autism. *Journal of Clinical Child & Adolescent Psychology*, *37*, 8-38. doi: 10.1080/15374410701817808.
- Rutter, M., Bailey, A., & Lord, C. (2003). *Social Communication Questionnaire (SCQ)*. Los Angeles: Western Psychological Services.

- Rutter, M., Le Couteur, A., & Lord, C. (2003). *Autism Diagnostic Interview Revised. Manual*. Los Angeles: Western Psychological Services.
- 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. doi: 10.1007/s10803-007-0407-z.
- Swinkels, S. H. N., Dietz, C., van Daalen, E., Kerkhof, I. H. G. M., van Engeland, H., & Buitelaar, J. K. (2006). Screening for autistic spectrum disorder in children aged 14 to 15 months. I: The development of the Early Screening of Autistic Traits questionnaire (ESAT). *Journal of Autism and Developmental Disorders*, *36*, 723-732. doi: 10.1007/s10803-006-0115-0.
- Trajman, A., & Luiz, R. R. (2008). McNemar  $\chi^2$  test revisited: comparing sensitivity and specificity of diagnostic examinations. *The Scandinavian Journal of Clinical & Laboratory Investigation*, *68*, 77-80. doi: 10.1080/00365510701666031.
- Vergara, I. A., Norambuena, T., Ferrada, E., Slater, A. W., & Melo, F. (2008). StAR: a simple tool for the statistical comparison of ROC curves. *BMC Bioinformatics*, *9*, 265-269. doi: 10.1186/1471-2105-9-265.
- Warreyn, P., Raymaekers, R., & Roeyers, H. (2004). *Handleiding Vragenlijst Sociale Communicatie*. Destelbergen: SIG vzw.
- Watson, L. R., Baranek, G. T., Crais, E. R., Reznick, J. S., Dykstra, J., & Perryman, T. (2007). The First Year Inventory: Retrospective parent responses to a questionnaire designed to identify one-year-olds at risk for autism. *Journal of Autism and Developmental Disorders*, *37*, 49-61. doi: 10.1007/s10803-006-0334-4.

Wetherby, A., & Prizant, B. (2002). *Communication and Symbolic Behavior Scales Developmental Profile. First normed edition.* . Baltimore, Maryland: Paul H Brookes Publishing.

Zink, I., & Lejaegere, M. (2003). *N-CDIs. Lijsten voor Communicatieve Ontwikkeling. Korte vormen.* Leuven/Leusden: Acco.

### Author Note

Mieke Dereu, Ruth Raymaekers, Petra Warreyn, Inge Schietecatte, Mieke Meirsschaut, and Herbert Roeyers, Research Group Developmental Disorders, Department of Experimental Clinical and Health Psychology, Ghent University, Belgium.

**Acknowledgments:** Partial funding for this research was provided by support from Steunpunt Expertisenetwerken and Vlaamse Vereniging Autisme.

We thank the day-care centres, the children and their families, without whom this study would not have been possible. Special thanks goes to the diagnostic centra across Flanders (COS Gent, COS Antwerpen, COS Brussel, COS Leuven, RCA Gent, RCA Antwerpen, RCA Brussel and ECA Leuven) for the good collaboration which made early diagnosis for screen positive children possible.

Correspondence should be addressed to Mieke Dereu, Ghent University, Department of Experimental Clinical and Health Psychology, Henri Dunantlaan 2, B-9000 Ghent, Belgium.

E-mail: [mieke.dereu@ugent.be](mailto:mieke.dereu@ugent.be).

Table 1

*Age and gender for the different screening instruments*

	CESDD	ESAT	M-CHAT	SCQ
	( <i>N</i> = 357)	( <i>N</i> = 352)	( <i>N</i> = 199)	( <i>N</i> = 152)
Mean age ( <i>SD</i> )	20.50 (5.78)	23.18 (6.05)	21.17 (2.57)	28.00 (5.00)
Age range	5.57 – 38.40	8.20 – 48.13	16.70 – 31.00	17.17 – 48.13
% Male	61.34	61.08	63.82	63.82

Table 2

*Age, gender, socio-economic status and test results for children seen at the university lab (n = 70)*

	ND			DD			sASD			ASD			Total		
	Mean	(SD)	n	Mean	(SD)	n	Mean	(SD)	n	Mean	(SD)	n	Mean	(SD)	n
Age in months	30.00	(5.46)	14	26.83	(6.15)	27	24.97	(6.98)	6	31.22	(8.10)	23	28.75	(7.02)	70
SES	51.64	(7.07)	14	47.06	(11.72)	27	40.17	(17.98)	6	42.61	(15.75)	23	45.92	(13.32)	70
MSEL															
Visual reception AE	32.86	(7.49)	14	24.07	(8.03)	27	25.17	(11.75)	6	27.17	(11.11)	23	26.94	(9.75)	70
Fine motor AE	30.14	(7.72)	14	22.81	(5.80)	27	20.83	(6.62)	6	25.22	(8.55)	23	24.90	(7.69)	70
Receptive language AE	30.29	(6.49)	14	19.52	(6.55)	27	22.00	(12.82)	6	23.17	(10.83)	23	23.09	(9.44)	70
Expressive language AE	30.00	(9.14)	14	17.22	(6.33)	27	19.83	(10.46)	6	19.96	(12.23)	23	20.09	(10.44)	70
Early Learning Composite	104.14	(14.19)	14	75.11	(16.49)	27	85.33	(22.22)	6	74.48	(21.40)	23	81.59	(21.45)	70
ADOS module 1 <sup>a</sup>															
Social domain	1.08	(1.31)	12	1.91	(2.21)	23	6.60	(2.30)	5	6.11	(3.11)	19	3.49	(3.29)	59
Communication domain	1.25	(1.22)	12	1.48	(1.56)	23	3.60	(1.34)	5	4.20	(2.31)	19	2.42	(2.14)	59
Repetitive domain	0.50	(0.90)	12	0.83	(0.83)	23	2.80	(1.92)	5	1.95	(1.51)	19	1.29	(1.39)	59
ADOS module 2															
Social domain	0.50	(0.71)	2	0.00	-	1	6.00	-	1	6.25	(1.71)	4	4.00	(3.25)	8
Communication domain	0.50	(0.71)	2	2.00	-	1	6.00	-	1	4.75	(0.50)	4	3.50	(2.20)	8
Repetitive domain	0.00	(0.00)	2	1.00	-	1	0.00	-	1	2.50	(1.73)	4	1.38	(1.69)	8
% Male	71.43 (n = 14)			66.67 (n = 27)			83.33 (n = 6)			73.91 (n = 23)			71.43 (n = 70)		

*ND* no diagnosis, *DD* developmental delay, *sASD* highly suspected of autism spectrum disorder, *ASD* autism spectrum disorder, *SES* socio-economic status (Hollingshead, 1975), *MSEL* Mullen Scales of Early Learning (Mullen, 1995), *AE* age equivalent, *ADOS* Autism Diagnostic Observation Schedule (Lord et al., 1999).

<sup>a</sup> The ADOS could not be administered in three children with DD because they could not yet move around freely in the test room.

Table 3

*Mean scores on the CESDD, ESAT, M-CHAT and SCQ for high- versus medium-risk children*

	<i>M</i>	<i>SD</i>	Range	Mann-Whitney <i>U</i>
CESDD				- <sup>a</sup>
High-risk ( <i>n</i> = 151)	5.92	4.07	2 - 20	
Medium-risk ( <i>n</i> = 206)	0.15	0.35	0 - 1	
Total ( <i>N</i> = 357)	2.59	3.90	0 - 20	
ESAT				11,258.00**
High-risk ( <i>n</i> = 147)	0.97	1.66	0 - 11	
Medium-risk ( <i>n</i> = 205)	0.30	0.68	0 - 5	
Total ( <i>N</i> = 352)	0.58	1.23	0 - 11	
M-CHAT				3,049.50**
High-risk ( <i>n</i> = 64)	2.34	3.08	0 - 11	
Medium-risk ( <i>n</i> = 135)	0.71	0.98	0 - 4	
Total ( <i>N</i> = 199)	1.24	2.06	0 - 11	
SCQ				1,965.50*
High-risk ( <i>n</i> = 62)	9.23	5.75	0 - 23	
Medium-risk ( <i>n</i> = 90)	6.33	3.75	0 - 21	
Total ( <i>N</i> = 152)	7.51	4.87	0 - 23	

<sup>a</sup> This value is not calculated because group membership was based on the CESDD scores.

\*  $p < .01$ ; \*\*  $p = .001$

Table 4

*Mann-Whitney U test for pairwise comparisons of total scores for different diagnostic groups*

	Mean score (SD)	ND	DD	sASD	ASD
<b>CESDD (N = 357)</b>					
ND (n = 289)	1.73 (2.90)		2,739.00***	138.50***	1,166.00***
DD (n = 35)	4.51 (4.43)			51.50*	297.00*
sASD (n = 7)	9.29 (5.47)				75.50
ASD (n = 26)	7.73 (5.66)				
<b>ESAT (N = 352)</b>					
ND (n = 289)	0.33 (0.63)		3,730.50*	374.50***	1,534.50***
DD (n = 33)	1.15 (1.89)			65.50*	267.00*
sASD (n = 7)	2.86 (2.34)				63.00
ASD (n = 23)	2.17 (2.61)				
<b>M-CHAT (N = 199)</b>					
ND (n = 172)	0.78 (1.11)		643.00**	39.00***	411.50**
DD (n = 13)	3.31 (3.52)			13.00*	54.00
sASD (n = 5)	7.40 (4.04)				10.00
ASD (n = 9)	3.56 (3.32)				
<b>SCQ (N = 152)</b>					
ND (n = 113)	5.98 (3.39)		444.50**	162.00	198.00***
DD (n = 16)	9.31 (3.50)			30.00	60.00**
sASD (n = 4)	9.00 (7.07)				18.50
ASD (n = 19)	14.79 (5.89)				

*ND* no diagnosis, *DD* developmental delay, *sASD* highly suspected of autism spectrum disorder, *ASD* autism spectrum disorder

\*  $p = .05$ ; \*  $p < .05$ ; \*\*  $p < .01$ ; \*\*\*  $p < .001$

Table 5

*Discriminant ability of the CESDD, ESAT and M-CHAT compared (n = 197)*

	CESDD		ESAT		M-CHAT	
Se (95% CI)	0.92	(0.62 – 1.00)	0.38	(0.15 – 0.68)	0.69	(0.39 – 0.90)
Sp (95% CI)	0.73	(0.66 – 0.79)	0.97	(0.93 – 0.99)	0.88	(0.82 – 0.92)
PPV (95% CI)	0.19	(0.11 – 0.32)	0.45	(0.18 – 0.75)	0.28	(0.14 – 0.47)
NPV (95% CI)	0.99	(0.95 – 1.00)	0.96	(0.91 – 0.98)	0.98	(0.94 – 0.99)
AUC (95% CI)	0.89	(0.83 – 0.96)	0.80	(0.66 – 0.95)	0.82	(0.68 – 0.97)

*Se* Sensitivity, *Sp* Specificity, *PPV* Positive Predictive Value, *NPV* Negative Predictive Value,

*AUC* Area Under the Curve

Table 6

*Discriminant ability of the CESDD, ESAT and SCQ compared (n = 149)*

	CESDD		ESAT		SCQ	
Se (95% CI)	0.90	(0.68 – 0.98)	0.38	(0.19 – 0.61)	0.71	(0.48 – 0.88)
Sp (95% CI)	0.68	(0.59 – 0.76)	0.98	(0.93 – 0.99)	0.85	(0.78 – 0.91)
PPV (95% CI)	0.32	(0.21 – 0.45)	0.73	(0.39 – 0.93)	0.44	(0.28 – 0.62)
NPV (95% CI)	0.98	(0.91 – 1.00)	0.91	(0.84 – 0.95)	0.95	(0.89 – 0.98)
AUC (95% CI)	0.87	(0.79 – 0.95)	0.75	(0.62 – 0.89)	0.85	(0.73 – 0.97)

*Se* Sensitivity, *Sp* Specificity, *PPV* Positive Predictive Value, *NPV* Negative Predictive Value,

*AUC* Area Under the Curve

Table 7

*Item analyses of the CESDD, ESAT, M-CHAT and SCQ*

	Non-ASD (%)	ASD (%)	Phi
<b>CESDD (N = 357<sup>a</sup>)</b>			
1. Lack of eye contact	7.41	27.27	0.20**
2. Abnormal eye contact	10.19	30.30	0.18*
3. Lack of response to name	6.17	21.21	0.17*
4. Seems to be deaf	1.85	12.12	0.18*
<b>5. Lack of social smile</b>	<b>8.95</b>	<b>45.45</b>	<b>0.32**</b>
6. Lack of exploration/passivity	12.35	27.27	0.13
7. Doesn't like being touched or cuddled	8.02	27.27	0.19**
8. Stereotyped behaviour	10.19	36.36	0.23**
9. Unusual sensory behaviour	3.70	24.24	0.26**
10. Lack of varied manipulative play	10.80	33.33	0.20**
11. Lack of shared enjoyment in social play	6.79	24.24	0.18*
<b>12. Lack of imitation</b>	<b>12.96</b>	<b>51.52</b>	<b>0.30**</b>
13. Lack of gestures	11.07	35.48	0.21**
14. Prefers to be alone	10.75	41.94	0.27**
<b>15. Lack of point or gaze following</b>	<b>3.91</b>	<b>35.48</b>	<b>0.36**</b>
16. Lack of pointing to indicate interest	12.70	48.39	0.28**
<b>17. Lack of showing</b>	<b>16.94</b>	<b>64.52</b>	<b>0.34**</b>
18. Easily frustrated	17.26	48.39	0.22**
19. Lack of facial expressions	8.79	25.81	0.16*
<b>20. Lack of functional play</b>	<b>10.10</b>	<b>45.16</b>	<b>0.30**</b>
21. Lack of anticipation of being picked up	6.51	22.58	0.17*
22. Unusual postures	6.51	29.03	0.23**
<b>23. Use of someone's hand</b>	<b>1.63</b>	<b>29.03</b>	<b>0.40**</b>
<b>24. Unusual interest or stereotyped play</b>	<b>5.63</b>	<b>36.84</b>	<b>0.39**</b>
<b>25. Lack of symbolic play</b>	<b>8.45</b>	<b>42.11</b>	<b>0.38**</b>
<b>ESAT (N = 352)</b>			
1. No interest in different toys	0.31	0.00	-0.02
<b>2. No varied play</b>	<b>4.04</b>	<b>33.33</b>	<b>0.33**</b>
3. Emotions not easily understandable	2.80	10.00	0.11
<b>4. Unusual reaction to sensory stimuli</b>	<b>0.31</b>	<b>13.33</b>	<b>0.31**</b>
5. No clear facial expressions	2.48	13.33	0.17*
<b>6. Not easy to make eye contact</b>	<b>2.80</b>	<b>30.00</b>	<b>0.35**</b>
7. Doesn't attract attention	13.04	26.67	0.11
8. Stereotypical movement	5.59	30.00	0.26**
<b>9. No showing</b>	<b>2.48</b>	<b>26.67</b>	<b>0.32**</b>
<b>10. No interest in other children or adults</b>	<b>0.31</b>	<b>16.67</b>	<b>0.35**</b>
11. Doesn't like cuddling	3.42	3.33	-0.00
12. No smile directed to others	1.24	6.67	0.12
13. Doesn't enjoy social play	0.62	10.00	0.22**
14. Doesn't react when spoken to	1.55	10.00	0.16*
<b>M-CHAT (N = 199)</b>			
1. Doesn't enjoy being swung	1.62	0.00	-0.03
2. No interest in other children	0.54	14.29	0.29**
<b>3. Doesn't like climbing on things</b>	<b>1.08</b>	<b>21.43</b>	<b>0.33**</b>
4. Doesn't enjoy social play	0.00	7.14	0.26**
5. No pretend play	3.78	28.57	0.28**
<b>6. No pointing to ask for something</b>	<b>1.62</b>	<b>35.71</b>	<b>0.44**</b>
<b>7. No pointing to indicate interest</b>	<b>3.78</b>	<b>42.86</b>	<b>0.40**</b>
8. No appropriate functional play	5.41	7.14	0.02
<b>9. No showing</b>	<b>1.62</b>	<b>28.57</b>	<b>0.37**</b>
10. No eye contact longer than 2 sec	1.62	14.29	0.21*
11. Oversensitive to noise	16.76	42.86	0.17
12. No social smile	0.54	14.29	0.29**

13. No imitation	8.11	21.43	0.12
14. No response to name	1.62	0.00	-0.03
15. No response to joint attention	3.78	28.57	0.28**
<b>16. Doesn't walk</b>	<b>3.24</b>	<b>35.71</b>	<b>0.36**</b>
17. No gaze following	3.78	21.43	0.21*
18. Unusual finger movements	3.78	28.57	0.28**
19. Doesn't attract attention to own activity	7.03	28.57	0.20*
20. Parents wondered if child was deaf	3.24	14.29	0.14
21. Doesn't understand what people say	3.24	21.43	0.22*
22. Stares or wanders	15.14	14.29	-0.01
23. No social referencing	4.32	21.43	0.19*
SCQ ( <i>N</i> = 152)			
2. No conversation	10.08	30.43	0.22*
3. Stereotyped utterances	52.71	39.13	-0.10
4. Inappropriate questions	10.08	13.04	0.04
5. Pronoun reversal	15.50	17.39	0.02
6. Neologisms	23.26	39.13	0.13
7. Verbal rituals	43.41	47.83	0.03
8. Compulsions and rituals	1.55	8.70	0.16
9. Inappropriate facial expression	22.48	26.09	0.03
10. Use of other's body	13.18	34.78	0.21
11. Unusual preoccupations	24.81	39.13	0.12
12. Repetitive use of objects	10.85	39.13	0.28**
13. Circumscribed interests	3.10	21.74	0.28**
<b>14. Unusual sensory interests</b>	<b>3.10</b>	<b>43.48</b>	<b>0.50**</b>
15. Hand and finger mannerisms	27.13	43.48	0.13
16. Complex body mannerisms	10.08	34.78	0.26*
17. Self injury	12.40	13.04	0.01
18. Unusual attachment to objects	6.20	4.35	-0.03
19. No friends	51.16	73.91	0.16
20. No social chat	41.09	43.48	0.02
21. No imitation	5.43	8.70	0.05
<b>22. No pointing to express interest</b>	<b>5.43</b>	<b>34.78</b>	<b>0.35**</b>
23. No gestures	43.41	60.87	0.13
24. No nodding to mean 'yes'	21.71	52.17	0.25*
25. No head shaking to mean 'no'	17.05	30.43	0.12
26. No eye gaze	5.43	21.74	0.22*
27. No social smiling	3.10	21.74	0.28**
28. No showing	3.88	13.04	0.15
29. Doesn't offer to share	13.18	34.78	0.21
30. Doesn't share enjoyment	4.65	17.39	0.18
31. Doesn't offer comfort	15.50	47.83	0.29**
32. Poor quality of social overtures	2.33	8.70	0.13
33. Limited range of facial expression	3.10	4.35	0.03
<b>34. No imitative social play</b>	<b>9.30</b>	<b>47.83</b>	<b>0.39**</b>
35. No imaginative play	21.71	56.52	0.28*
<b>36. No interest in children</b>	<b>7.75</b>	<b>47.83</b>	<b>0.42**</b>
37. No response to other children	10.85	21.74	0.12
<b>38. No attention to voice</b>	<b>8.53</b>	<b>47.83</b>	<b>0.40**</b>
<b>39. No imaginative play with peers</b>	<b>27.91</b>	<b>78.26</b>	<b>0.38**</b>
<b>40. No group play</b>	<b>22.48</b>	<b>60.87</b>	<b>0.31**</b>

Items with Phi values of .30 or above are printed in bold.

<sup>a</sup> For the CESDD, items 13 to 23 were completed for 338 children above 12 months and items 24 and 25 were filled out for 90 children above 24 months

\*  $p < .01$ ; \*\*  $p < .001$

## Figure Captions

*Figure 1.* Overall screening design

*Figure 2.* Return rates for the parent questionnaires

*Figure 3.* Parental compliance with the invitation for further assessment





