Title page with all author contact information; Abstract with key words and corresponding author email

Piloting the use of a short observation list for ASD-symptoms in day-care; challenges and further possibilities

Kenneth Larsen[1], Astrid Aasland[2], Synnve Schjølberg[4], Ulla I. Hansen[1] & Trond H. Diseth[2,3]

[1]Division of Pediatric and Adolescent Medicine, Regional Resource Centre for Autism, ADHD, Tourette's Syndrome and Narcolepsy, Oslo University Hospital, Oslo, Norway

[2]Department of Child and Adolescent Mental Health in Hospitals, Division of Paediatric and Adolescent Medicine, Oslo University Hospital, Oslo, Norway

[3]Faculty of Medicine, Institute of Clinical Medicine, University of Oslo, Oslo, Norway

[4]Norwegian Institute of Public Health, Oslo, Norway

Author Note

Corresponding author: Kenneth Larsen (k_larsen@mac.com)

Abstract

Early symptoms of Autism Spectrum Disorder (ASD) develop through the second year of life, making a stable ASD diagnosis possible around 24 months of age. However, in general, children with ASD are diagnosed later. In this study we explored the use of a short observation list to detect symptoms associated with ASD in children 12 to 24 months of age attending typical day-care centers. The results indicate that a short observation list used by day-care teachers does not reveal sufficient properties to be independently used in young children in day-care centers. Further studies should explore multiple and repeated measures for early detection of symptoms associated with ASD in typical day-care centers.

Keywords: autism spectrum disorder, early identification, day-care centers, screening

Abstract

Early symptoms of Autism Spectrum Disorder (ASD) develop through the second year of life, making a stable ASD diagnosis possible around 24 months of age. However, in general, children with ASD are diagnosed later. In this study, we explored the use of a short observation list to detect symptoms associated with ASD in children 12 to 24 months of age attending typical day-care centers. The results indicate that a short observation list used by day-care teachers does not reveal sufficient properties to be independently used in young children in day-care centers. Further studies should explore multiple and repeated measures for early detection of symptoms associated with ASD in typical day-care centers.

Keywords: autism spectrum disorder, early identification, day-care centers, screening

Compliance with Ethical Standards

All authors declare that they have no conflict of interests. And all participants included in this study have given informed consent, through parental consent.

Piloting the use of a short observation list for ASD-symptoms in day-care; challenges and further possibilities

Autism spectrum disorder (ASD) is characterized by persistent deficits in social communication and social interaction as well as restricted, repetitive patterns of behavior, interests, or activities, starting from early development (American Psychiatric Association, 2013). ASD is associated with significant impairment in daily functioning and is highly correlated with various comorbidities, including impairments in attention, mood, cognitive skills and adaptive skills (Jones, Gliga, Bedford, Charman, & Johnson, 2014). The prevalence of ASD in Norwegian children is found to be approximately 0.6% (Isaksen, Diseth, Schjølberg, & Skjeldal, 2012; Suren et al., 2012). The worldwide prevalence estimate is approximately 1% (Elsabbagh et al., 2012).

The universal etiology of autism is unclear, making identification of ASD relying on behavioral markers and developmental trajectories (Lord, Elsabbagh, Baird, & Veenstra-Vanderweele, 2018; Zwaigenbaum & Penner, 2018). The mean age of diagnosis for ASD is reported to be approximately four years (Larsen, 2015; Mazurek et al., 2014). Recent research, especially research following high-risk infants, has revealed more detailed information on developmental trajectories associated with ASD, that may facilitate earlier diagnosis (Bailey, Palferman, Heavey, & Le Couteur, 1998; Bedford et al., 2012; Elsabbagh & Johnson, 2010; Gamliel, Yirmiya, & Sigman, 2007; Ozonoff et al., 2014; Shephard et al., 2018). Several studies have identified a characteristic development of social deficit emerging late in the first year of life and through the second year of life in children with ASD (Ozonoff et al., 2010; Richards, Mossey, & Robins, 2016). In addition, retrospective reports of symptoms by parents and others indicate that in early developmental age, the negative symptoms, namely deficits in social

interaction, are most frequently reported (Goodwin, Matthews, & Smith, 2018; Larsen, Aasland, & Diseth, 2018b). Early symptoms of ASD are related to both the diagnostic areas (APA, 2013) as well as other developmental areas (Gammer et al., 2015; Jones et al., 2014). The early symptoms of ASD affect early social communication and interaction (Gammer et al., 2015; Landa, Holman, & Garrett-Mayer, 2007; Sally Ozonoff et al., 2010; Rozga et al., 2011), restricted and stereotypical patterns of behavior (S. Ozonoff et al., 2008), delayed language development (Lazenby et al., 2015; Mitchell et al., 2006; Zwaigenbaum et al., 2005), problems with attentional control (M. Elsabbagh et al., 2013; Gammer et al., 2015), emotional regulation (Feldman, Hendry, Ward, Hudson, & Liu, 2014; Zwaigenbaum et al., 2005), and motor development (G. T. Baranek, 1999; Iverson & Wozniak, 2007; Leonard et al., 2013).

Brian, Bryson, and Zwaigenbaum (2015) have identified four key developmental domains that are predictive of ASD in early development (6 – 24 months) and thus represent key intervention targets. The identified key areas are 1) early attentional control, 2) emotional regulation, 3) social orientation/approach, and 4) communication development. To ensure sufficient scope in early identification of ASD and lay the foundation for planning early interventions in educational and family contexts, these four development areas may be important.

This new, emerging knowledge on early signs of ASD could result in more children being identified, assessed and diagnosed with ASD at the age of 12 and 24 months. Research has shown that ASD diagnoses made in the second year of life are stable and confirmed later in development (Ozonoff et al., 2015; Pierce et al., 2019; Zwaigenbaum et al., 2016). Identification of ASD early in development has the potential for reduced symptom burden, early intervention and improved quality of life for the children and their families (Hampton & Kaiser, 2016; Landa, 2018; Rotholz, Kinsman, Lacy, & Charles, 2017).

Several factors have been identified as associated with the age of diagnosis for children with ASD, such as level of autism symptoms, IQ and birth order, as well as socioeconomic factors (Fountain, King, & Bearman, 2011; Mazurek et al., 2014). Barriers to early identification may include factors as lack of the professionals' education, low level of parent education, professionals' hesitancy in giving early diagnosis, waiting-lists, economic factors in families and community, geographic location, and parent fear of stigma (Elder, Brasher, & Alexander, 2016; Khowaja, Hazzard, & Robins, 2015; Rosenberg, Landa, Law, Stuart, & Law, 2011).

Universal screening may enhance earlier identification of children at risk. However, there are controversies related to the need for and the usefulness of universal screening for ASD (Broder-Fingert, Feinberg, & Silverstein, 2018; Johnson & Myers, 2007; McPheeters et al., 2016; Øien et al., 2018). The primary goal of screening instruments is to enhance the effectiveness of early identification and enable implementation of early intervention for children with ASD (McPheeters et al., 2016), and studies have shown that screening leads to earlier referral and diagnosis (Miller et al., 2011). To facilitate widespread use, screening tools should be brief and be easy to implement. They also need to have high sensitivity to discover a high proportion of children with the disorder, as well as high specificity to separate children with ASD from children with other disorders (Cadman, Chambers, Feldman, & Sackett, 1984; Nah, Young, & Brewer, 2018). A recent review of current screening instruments (Marlow, Servili, & Tomlinson, 2019) identified 16 screening tools for ASD in the age below 24 months. Although most of the identified tools had sufficient sensitivity and specificity (>70 %), there are several challenges with current screening instruments for young children with ASD. Results from the Norwegian Mother, Father and Child Cohort Study (MoBa) indicate that screening based on parental questionnaires in large prospective cohorts may have low sensitivity and Positive

Predictive Value (Stenberg et al., 2014), and include a large proportion of false negatives (Øien et al., 2018). This is also documented in other studies (Chlebowski, Robins, Barton, & Fein, 2013; Guthrie et al., 2019; McPheeters et al., 2016; Robins et al., 2014; Stenberg et al., 2014). The cultural aspects of screening tools are not fully understood, and to date, there are no screening tools developed for use in Norway (Al Maskari, Melville, & Willis, 2018; Marlow et al., 2019).

Accurate screening is dependent on the quality of reporting. Screening for ASD is often carried out by physicians or staff at community health care centers through parental completion of a symptom checklist. Parents may under-report symptoms if they do not already suspect that their child has ASD (Havdahl et al., 2017). Input from other health care workers and teachers may increase the likelihood for early identification (Constantino et al., 2007; DeVincent, Gadow, Strong, Schwartz, & Cuva, 2008; Janvier et al., 2016; Larsen, Aasland, & Diseth, 2018a; Self, Coufal, & Parham, 2010). However, professionals in the health care system do not necessarily recognize concerns associated with ASD due to the limited time frame of health check-ups (Coury, 2015; Gabrielsen et al., 2015). Overall, young children aged 15-33 months with ASD display more typical than atypical behavior (Gabrielsen et al., 2015), indicating that another context that facilitates repeated and long-term observations may be needed to detect atypical behavior (Branson, Vigil, & Bingham, 2008; Dereu et al., 2012). Research indicates that long-term surveillance involving a parent-professional partnership may facilitate earlier identification of ASD in young children (Charman & Gotham, 2013; Duvekot, van der Ende, Verhulst, & Greaves-Lord, 2015; Janvier et al., 2016; Williams, 2016). Nilsson Jobs, Bölte, & Falck-Ytter (2018) found that parents and preschool staff rated symptoms associated with ASD with a high agreement at three years of age, indicating that day-care teachers' competence in typical development may be utilized to recognize deviant development. Hence, day-care centers and

kindergartens may be a suitable and effective arena for detection of ASD in a parent-professional partnership. Moreover, symptoms of ASD may be more visible in day-care settings as the focus is on play and social interaction and children spend many hours a day here. [redacted] found that day-care professionals and parents clearly identified symptoms that retrospectively distinguished between children with ASD and children with typical development between 12 and 24 months of age. Based on parents and day-care professionals´ rating of 73 early symptoms known to be associated with ASD, six defining symptoms were found to clearly separate between children with ASD and typically developing children. These six symptoms constitute the short observation list further explored in the present study. To our knowledge, there has not been developed an observation tool for use in day-care centers in the second year of life with the purpose to facilitate early identification of ASD.

The present study is the second part in developing and exploring the feasibility and psychometric properties of a short observation list (short-list), consisting of six items, to assess early symptoms associated with ASD in young children, aged 12 to 24 months, in Norwegian day-care centers. The development of the six-item checklist is described in [redacted]. The aim of the present study was to explore: 1. The distribution of day-care teachers' scores on the short-list; 2. The association between the day-care teachers' scores on the short-list and the assessment of symptoms on a standardized diagnostic instrument (ADOS-2) by a trained clinician; 3. Agreement of ratings between parents and day-care teachers on the short-list.

Methods

Participants

The participants are children, 12-24 months of age, attending regular day-care centers. The day-care centers were open to all children regardless of any developmental concerns and did

not include any special units. Children were included in the study based on parental consent. Parents were given information on all sides of the project in a meeting held in each day-care center. After the meeting, all parents with children 12-24 months of age were given written information explaining the details of the project, along with a consent form that was returned to the day-care center. From the 32 first day-care centers visited, parents of 139 children agreed for their child to participate. After written consent was received, parents and daycare teachers were asked to complete an online version of the short-list [redacted]. Of the 139 giving consent to participate, the day-care teachers provided data on 131 children. According to the estimated number of possible participants (those whose parents received the information), a total of 131 participants gives a response rate of approximately 51 %.

Based on day-care teachers' rating on the short-list, children were allocated into three groups: 1) children with no rating of behaviors associated with ASD (n= 72), 2) children with rating on one item on the short-list (n=29), and 3) children with rating on two or more items on the short-list (n=30). The groups were chosen based on the results from the retrospective study, suggesting a score on two items on the short-list as the cut-off indicating a heightened risk of ASD (Larsen et al., 2018b).

The intention was to include more day-care centers until we had at least 60 children in each of the three groups. However, the inclusion was stopped as the proportion of children who scored according to group 2 and 3 criteria was much higher than expected. Scoring in groups 2 and 3 indicated based on the retrospective study higher risk for ASD. These high proportions indicated that the short-list did not separate as well between children at risk and children not at risk as believed based on the retrospective study. Hence, we found it not ethically justified to include more children to assessment with ADOS-2, based on the current protocol.

All children in group 2 and 3, and 30 randomly selected children from group 1, a total of 89 children, were invited to an assessment with Autism Diagnostic Observation Schedule 2 (ADOS-2), toddler module (Lord, Luyster, Gotham, & Guthrie, 2012). A total of 68 of the 89 children invited were assessed with ADOS-2, giving a response rate of 76.4%. Reasons for declining the invitation to ADOS-2 assessment were illness, not consenting further participation and difficulties finding time to participate. There were no significant differences in scores on the short-list between the children assessed with ADOS-2, and the children invited, but not assessed.

Demographic information on all participants (n=131), participants assessed with ADOS-2 (n=68) and participants assessed with ADOS-2 in each of the three groups (Group 1: n=24, Group 2: n=20, and Group 3: n=24) are shown in table 1. All demographic information was provided through a parent questionnaire. No significant differences were found between the three groups, and no significant differences were found between the group assessed with ADOS-2 (n = 68) and the group not assessed with ADOS-2 (n = 63).

- Table 1 -

Measures

Short observation list (short-list)

The shortlist was developed in a previous study asking parents and day-care teaches to retrospectively rate children with an ASD-diagnosis and control children (3-6 years of age) on whether the children had displayed behavior associated with ASD or not at 12-24 months of age [redacted]. Parents and day-care teachers rated the children on 73 symptoms selected from published screening tools for ASD and from the literature on early symptoms associated with ASD in the second year of life. Based on this rating, the list was reduced to six symptoms

that discriminated excellent between ASD and children with typical development. These symptoms are listed in table 2. These six symptoms were rated as 0 = "not present" (the behavior did not occur), 1 = "partly present" (the behavior occurred sometimes or partly), or 2 = "definitely present" (the behavior was definitely present). In the present study, parents and day-care teachers were asked to rate how children currently performed compared to other children at the same age. The day-care teachers completing the short-list were required to have interacted with the child for at least two weeks. If they were uncertain about some items, they were instructed to observe for at least two more days before answering. Day-care teachers were asked to complete the short-list twice times with a two-week interval.

The short-list was scored both on the number of endorsed items (range 0-6), and total summary score (range 0-12). Two endorsed items (scored 1 or 2) was used as the cut-off in this study [redacted].

- Table 2 -

ADOS-2

The Autism Diagnostic Observation Schedule, second edition (ADOS-2) (Lord et al., 2012) is a comprehensive diagnostic instrument and is considered a part of the recommended "gold standard" for the assessment of ASD (Risi et al., 2006; Shumway et al., 2012). The recent revision of the ADOS-2 includes the addition of the Toddler Module (Lord et al., 2012) which can be used to assess symptoms associated with ASD in children 12-30 months old. The Toddler Module in ADOS-2 is designed to create loosely structured activities involving highly motivating materials and assess how and whether the child makes social initiations in order to maintain the interaction. Based on the activities in ADOS-2 the clinician codes distinct aspects of behavior. Each aspect of behavior is coded on a dimension from 0 (no evidence of abnormality) to 3

(markedly abnormal), all defined by specific descriptions of the behavior shown. Based on age and language level two different algorithms are used to compare each child's score with a predefined cut-off (Lord et al., 2012). The Toddler Module in ADOS-2 has shown good sensitivity and specificity for discriminating young children with ASD from children with typical development (Luyster et al., 2009).

ADOS-2 was in this study administrated by a research trained clinician (UIH) blind of the day-care teachers' rating on the short observation list.

Statistical analysis

Descriptive statistics were computed for both the total group and the group assessed by ADOS-2. Potential differences between groups were explored using t-tests and chi-square tests, for comparisons of several groups Kruskal-Wallis chi-squared was applied. Spearman rank correlations were used to compare the ratings on the short observation list and ADOS-2. Both comparisons of the total summary scores and scores on individual items that logically may be related were explored. The agreement between day-care professionals and parents was examined by the intraclass correlation coefficient (ICC). All statistical analyses were done in R version 3.5.2.

This study was approved by the Regional Committee for Medical and Health Research Ethics (ref.no 2017/481).

Results

The distribution of the day-care teachers' total summary scores on the short-list in the total group of 131 children is shown in figure 1. The mean total summary score was 1.19, SD 1.88, range (0-9). The mean number of items endorsed on the short-list in the total group was

0.88, SD 1.23, range (0-6). The number of children in the total group having a rating (of 1 or 2) on each item is shown in figure 2. One hundred and one children were rated twice. The agreement between the day-care professionals´ two ratings of the total summary score was high (ICC 0.85 (0.79-0.89). Further day-care teachers indicating previous experience with ASD rate slightly fewer symptoms (0.67) than day-care teachers who do not (1.01), however, this difference was not significant (p=0.263).

When using the cut-off proposed by the retrospective study (a score of 1 or 2 on two or more items) (Larsen et al., 2018b), 22.9 % (n=30) of the children in the total group scored above cut-off and additionally 22.1% of the children (n=29) scored in the borderline area (a score of 1 or 2 on one item).

- Figure 1 here –
- Figure 2 here –
- Figure 3 here -

Scores on ADOS-2 were then computed for each of the three groups, see table 2. The mean ADOS-2 scores were not significantly different across the three groups (Kruskal-Wallis chi-squared = 4.57(2), p = 0.102), although there was a trend for children in group 3 having higher scores on ADOS-2 than children in the two other groups. This trend was mainly due to the large variability in scores within each group, especially in group 3.

- Table 3 -

There was no significant correlation between the total summary score on the short observation list and the total ADOS-2 score (rho = 0.19, p = 0.126), or the number of items rated 1 or 2 on the short observation list and the total ADOS-2 score (rho = 0.23, p = 0.058) although there was a slight trend. Nevertheless, looking at individual items, there was a significant

correlation between day-care teachers' rating of "*Response to joint attention*" (rho = 0.27, p = 0.024), and rating on "*Initiates joint attention*" (rho = 0.259, p = 0.032) with the total ADOS-2 score. Further analysis also showed that the item "*Response to joint attention*" was significantly associated with the ADOS-2 items "*Shared enjoyment in interaction*" (rho = 0.282, p = 0.02), "*Amount of social input/maintenance of attention*": Examiner (rho = 0.364, p > 0.01), and "*Functional and symbolic imitation*" (rho = 0.239, p = 0.05). The item "*Initiates joint attention*" on the short-list was significantly correlated with the ADOS-2 items "*Functional and symbolic imitation*" (rho = 0.259, p = 0.03), and "*Anxiety*" (rho = 0.260, p = 0.03).

Intraclass correlation coefficient for rating by day-care teachers and parents for the full short observation list was .0449 (-.138 - .225). For individual items see table 4.

- Table 4 –

Discussion

This study explores if a short observation list could be useful for day-care teachers in identifying symptoms associated with ASD in children 12 to 24 months of age. The short-list shown to have good properties when used retrospectively in children already diagnosed with ASD [redacted], showed poorer properties when used with a random sample of young children in day-care centers. When using the cut-off proposed by the retrospective study (Larsen et al., 2018b), the use of the short-list suggested a prevalence of behavior suspicious of ASD in 23% - 45% of the children in this population. This study has a low number of participants, but based on a prevalence of 1 %, we could expect 1-2 children with ASD among the participants (Elsabbagh et al., 2012). The high proportion of children rated with behaviors associated with ASD prompted the termination of the inclusion of participants. The design of this study poses the

risk of self-selection bias, which may have influenced the results, in both initial recruitment and agreeing to ADOS-2 assessment, however, no significant differences were found between groups on the latter. Although the results displayed a trend to a correlation between the number of items rated on the short-list and ADOS-2 score, the high number of children scoring on or above cut-off indicated that the short-list did not have the psychometric properties needed to discriminate children with possible ASD and typically developing children in the day-care centers.

The results indicate that a relatively large number of children at this age display the behavior rated by the items included in the short-list. The age period between 12 and 24 months is characterized by rapid development in social and communication skills, and there is a large variability in this development (Gabrielsen et al., 2015; Morales et al., 2000; Reilly et al., 2009). Also, there is great variation in the normal development of language and attention in this age group. Due to this variability, the behaviors included in the short-list may be quite common as a variation of the normal development in this age. Hence, the age 12-24 months may be too early to detect deficits in these developmental areas by screening a population-based sample of children, also indicated by group 3 consisting of the youngest children.

The challenge to precisely score these behaviors in day-care children attending a group of children in everyday activities may also have impeded the specificity of the short-list. The wording of individual items may influence day-care teachers´ understanding and subsequent rating of behaviors. Further, the inclusion of more items exploring each behavior may be important to detect deviancies more precisely. The day-care teachers in the present study had received no education or training in identifying symptoms of ASD before completing the short-list. However, knowledge and experience with ASD may be an important factor in early detection of ASD (Ben-Sasson, Atun-Einy, Yahav-Jonas, Lev-On, & Gev, 2018; Constantino et

al., 2007) In the present study, day-care teachers who had previous experience with ASD, tended to rate fewer symptoms than day-care teachers without the same experience. The precise definition and quality of each item describing deviant development associated with ASD may be clearer to teachers experienced with ASD. A short training for teachers defining the terms in each item, combined with examples of deviant and normal development, may have improved the properties of the short-list (Ben-Sasson et al., 2018), simultaneously reducing the cost-effectiveness and efficiency of the short-list.

The children in the present study were rated twice during a two-week period. Rating only within a short timeframe may render the results vulnerable to periodical variations in behaviors. Inclusion of repeated points of measurement over a longer time-period may have improved the results of the short-list (Branson et al., 2008). Likewise, the inclusion of short structured social situations to facilitate observation of specific social behaviors may have improved the observation tool, as previously suggested by others (Bryson, Zwaigenbaum, McDermott, Rombough, & Brian, 2008). Further investigations should be done into tools and systems to support day-care teachers in detecting early signs of ASD.

The study measured the properties of the short-list against ADOS-2, as ADOS-2 is considered a part of the “gold standard” in the assessment of symptoms associated with ASD. Dividing the children into three groups based on the proposed cut-off [redacted], the assessment with ADOS-2 did not reveal any significant difference in ADOS-2 score between the groups. Even if the classification made by the short-list was not confirmed by a significant difference in ADOS-2 score, the number of items rated on the short- list showed a trend towards being associated with the ADOS-2 score. This finding suggests that the items included in the short-list may be relevant in observing and assessing early social and communicative

development. The items assessing joint attention in the short-list were those with the strongest correlation with the total ADOS-2 scores, with "*Response to joint attention*" and "*Initiating joint attention*" showing a significant correlation with the total ADOS 2-score. Even if "*Response to joint attention*" was correlated to the total ADOS-2 score, the correlation did not reach significance with the specific item "*Response to joint attention*" in ADOS-2. It was significantly correlated with the ADOS-2 items "*Shared engagement in interaction*" and "*Functional and symbolic imitation*". Responses to joint attention are early developmental skills central to language and social development and deficits may be indicative of development associated with later ASD (Dawson et al., 2004). Our results may support that assessment of deficits in joint attention in the age between 12-24 months of age in day-care centers may be of value in the early identification of ASD, as previously suggested by others (Goodwin et al., 2018; Mundy, Sigman, & Kasari, 1994; Ozonoff et al., 2010). However, the lack of a significant correlation between the item on the short-list and the ADOS-2 should prompt further investigations on how to best assess joint attention in a typical day-care setting. The age between 12 and 24 months may give challenges in selecting the correct ADOS-2 module. Children usually start speaking in sentences in this age period, which prompts the use of module 2. The choice of module may impact the score on the items included in ADOS-2 (Klein-Tasman, Risi, & Lord, 2007). However, the ADOS-2 was administrated by a research-trained professional that also evaluated the appropriateness of the module.

This study indicates low agreement between parents and day-care professionals in observing children not identified with developmental challenges according to the short-list. When used retrospectively, the same observation list indicated excellent agreement (Larsen et al., 2018a). This is in line with our other findings indicating that the items included in the

observation list need further specification and that knowledge on behaviors typically observed in early autism may be necessary to observe these symptoms.

This study experience some common challenges with screening for ASD (Broder-Fingert et al., 2018; Øien et al., 2018; Stenberg et al., 2014). The need for time-efficient and easy to implement screeners compromise the sensitivity, specificity and positive predictive value of the screening tool. The risk of large numbers of false positives may lead to unnecessary concern and a possible change in the perception of children without any basis in real developmental concerns. A high proportion of false positives compromise the effectiveness of the diagnostic process and increase the load on already oversubscribed clinical recourses (Guthrie et al., 2019; Stenberg et al., 2014). Further, low specificity of a screener or observation tool may lead to numerous false negatives (Øien et al., 2018). These cases may effectively be identified later in development. However, a negative screen may delay identification of concern and diagnosis.

Based on the results of this study, the proposed short-list does not show qualities that support the use in mainstream day-care centers. In the effort to increase the number of children with ASD receiving an early diagnosis, we find it important also to report less successful results. These results may be important for and guide further research efforts to develop more effective tools and systems for early detection of ASD. Further studies of early detection of ASD in typical day-care settings could explore whether training of day-care teachers and repeated measures over a longer time span could improve the psychometric properties of a screener. The possibility and effect of multiple sources of information could also be further explored in typical day-care settings.

References

Al Maskari, T. S., Melville, C. A., & Willis, D. S. (2018). Systematic review: Cultural adaptation and feasibility of screening for autism in non-English speaking countries. *International Journal of Mental Health Systems*, *12*(1). https://doi.org/10.1186/s13033-018-0200-8

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub.

Bailey, A., Palferman, S., Heavey, L., & Le Couteur, A. (1998). Autism: The phenotype in relatives. *Journal of Autism and Developmental Disorders*, *28*(5), 369–392.

Bedford, R., Elsabbagh, M., Gliga, T., Pickles, A., Senju, A., Charman, T., & Johnson, M. H. (2012). Precursors to social and communication difficulties in infants at-risk for autism: Gaze following and attentional engagement. *Journal of Autism and Developmental Disorders*, *42*(10), 2208–2218.

Ben-Sasson, A., Atun-Einy, O., Yahav-Jonas, G., Lev-On, S., & Gev, T. (2018). Training Physical Therapists in Early ASD Screening. *Journal of Autism and Developmental Disorders*, 1–13.

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*(6), 523–530.

Broder-Fingert, S., Feinberg, E., & Silverstein, M. (2018). Improving Screening for Autism Spectrum Disorder: Is It Time for Something New? *Pediatrics*, e20180965.

Bryson, S. E., Zwaigenbaum, L., McDermott, C., Rombough, V., & Brian, J. (2008). The

Autism Observation Scale for Infants: Scale development and reliability data. *Journal of Autism and Developmental Disorders*, *38*(4), 731–738.

Cadman D, Chambers L, Feldman W, & Sackett D. (1984). Assessing the effectiveness of community screening programs. *JAMA*, *251*(12), 1580–1585. https://doi.org/10.1001/jama.1984.03340360046028

Charman, T., & Gotham, K. (2013). Measurement Issues: Screening and diagnostic instruments for autism spectrum disorders–lessons from research and practise. *Child and Adolescent Mental Health*, *18*(1), 52–63.

Chlebowski, C., Robins, D. L., Barton, M. L., & Fein, D. (2013). Large-scale use of the modified checklist for autism in low-risk toddlers. *Pediatrics*, peds-2012.

Constantino, J. N., Lavesser, P. D., Zhang, Y. I., Abbacchi, A. M., Gray, T., & Todd, R. D. (2007). Rapid quantitative assessment of autistic social impairment by classroom teachers. *Journal of the American Academy of Child & Adolescent Psychiatry*, *46*(12), 1668–1676.

Coury, D. L. (2015). Babies, bathwater, and screening for autism spectrum disorder: Comments on the USPSTF recommendations for autism spectrum disorder screening. *Journal of Developmental & Behavioral Pediatrics*, *36*(9), 661–663.

Dawson, G., Toth, K., Abbott, R., Osterling, J., Munson, J., Estes, A., & Liaw, J. (2004). Early Social Attention Impairments in Autism: Social Orienting, Joint Attention, and Attention to Distress. *Developmental Psychology*, *40*(2), 271–283. https://doi.org/10.1037/0012-1649.40.2.271

Dereu, M., Raymaekers, R., Warreyn, P., Schietecatte, I., Meirsschaut, M., & Roeyers, H. (2012). 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.

Journal of Autism and Developmental Disorders, *42*(5), 781–796.

DeVincent, C., Gadow, K. D., Strong, G., Schwartz, J., & Cuva, S. (2008). Screening for autism spectrum disorder with the Early Childhood Inventory-4. *Journal of Developmental and Behavioral Pediatrics : JDBP*, *29*(1), 1–10. https://doi.org/10.1097/DBP.0b013e3181468c32

Duvekot, J., van der Ende, J., Verhulst, F. C., & Greaves-Lord, K. (2015). The screening accuracy of the parent and teacher-reported Social Responsiveness Scale (SRS): Comparison with the 3Di and ADOS. *Journal of Autism and Developmental Disorders*, *45*(6), 1658–1672.

Elder, J. H., Brasher, S., & Alexander, B. (2016). Identifying the Barriers to Early Diagnosis and Treatment in Underserved Individuals with Autism Spectrum Disorders (ASD) and Their Families: A Qualitative Study. *Issues in Mental Health Nursing*, *37*(6), 412–420. https://doi.org/10.3109/01612840.2016.1153174

Elsabbagh, M., Divan, G., Koh, Y.-J., Kim, Y. S., Kauchali, S., Marcín, C., … Wang, C. (2012). Global prevalence of autism and other pervasive developmental disorders. *Autism Research*, *5*(3), 160–179.

Elsabbagh, M., & Johnson, M. H. (2010). Getting answers from babies about autism. *Trends in Cognitive Sciences*, *14*(2), 81–87.

Fountain, C., King, M. D., & Bearman, P. S. (2011). Age of diagnosis for autism: Individual and community factors across 10 birth cohorts. *Journal of Epidemiology & Community Health*, *65*(6), 503–510.

Gabrielsen, T. P., Farley, M., Speer, L., Villalobos, M., Baker, C. N., & Miller, J. (2015). Identifying autism in a brief observation. *Pediatrics*, peds-2014.

Gamliel, I., Yirmiya, N., & Sigman, M. (2007). The development of young siblings of children with autism from 4 to 54 months. *Journal of Autism and Developmental Disorders*,

37(1), 171–183.

Goodwin, A., Matthews, N. L., & Smith, C. J. (2018). Parent-reported early symptoms of autism spectrum disorder in children without intellectual disability who were diagnosed at school age. *Autism*, 1362361318777243. https://doi.org/10.1177/1362361318777243

Guthrie, W., Wallis, K., Bennett, A., Brooks, E., Dudley, J., Gerdes, M., … Miller, J. S. (2019). Accuracy of Autism Screening in a Large Pediatric Network. *Pediatrics*, *144*(4), e20183963. https://doi.org/10.1542/peds.2018-3963

Hampton, L. H., & Kaiser, A. P. (2016). Intervention effects on spoken□language outcomes for children with autism: A systematic review and meta□analysis. *Journal of Intellectual Disability Research*, *60*(5), 444–463.

Havdahl, K. A., Bishop, S. L., Surén, P., Øyen, A., Lord, C., Pickles, A., … Hornig, M. (2017). The influence of parental concern on the utility of autism diagnostic instruments. *Autism Research*, *10*(10), 1672–1686.

Isaksen, J., Diseth, T. H., Schjølberg, S., & Skjeldal, O. H. (2012). Observed prevalence of autism spectrum disorders in two Norwegian counties. *European Journal of Paediatric Neurology*, *16*(6), 592–598.

Janvier, Y. M., Harris, J. F., Coffield, C. N., Louis, B., Xie, M., Cidav, Z., & Mandell, D. S. (2016). Screening for autism spectrum disorder in underserved communities: Early childcare providers as reporters. *Autism*, *20*(3), 364–373.

Johnson, C. P., & Myers, S. M. (2007). Identification and evaluation of children with autism spectrum disorders. *Pediatrics*, *120*(5), 1183–1215.

Jones, E. J., Gliga, T., Bedford, R., Charman, T., & Johnson, M. H. (2014). Developmental pathways to autism: A review of prospective studies of infants at risk.

Neuroscience & Biobehavioral Reviews, *39*, 1–33.

Khowaja, M. K., Hazzard, A. P., & Robins, D. L. (2015). Sociodemographic barriers to early detection of autism: Screening and evaluation using the M-CHAT, M-CHAT-R, and follow-up. *Journal of Autism and Developmental Disorders*, *45*(6), 1797–1808.

Klein-Tasman, B. P., Risi, S., & Lord, C. E. (2007). Effect of Language and Task Demands on the Diagnostic Effectiveness of the Autism Diagnostic Observation Schedule: The Impact of Module Choice. *Journal of Autism and Developmental Disorders*, *37*(7), 1224–1234. https://doi.org/10.1007/s10803-006-0266-z

Landa, R. J. (2018). Efficacy of early interventions for infants and young children with, and at risk for, autism spectrum disorders. *International Review of Psychiatry*, *30*(1), 25–39. https://doi.org/10.1080/09540261.2018.1432574

Larsen, K. (2015). The early diagnosis of preschool children with autism spectrum disorder in Norway: A study of diagnostic age and its associated factors. *Scandinavian Journal of Child and Adolescent Psychiatry and Psychology*, *3*(2), 136–145.

Larsen, K., Aasland, A., & Diseth, T. H. (2018a). Brief report: Agreement between parents and day-care professionals on early symptoms associated with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *48*(4), 1063–1068.

Larsen, K., Aasland, A., & Diseth, T. H. (2018b). Identification of Symptoms of Autism Spectrum Disorders in the Second Year of Life at Day-Care Centres by Day-Care Staff: Step One in the Development of a Short Observation List. *Journal of Autism and Developmental Disorders*, *48*(7), 2267–2277.

Lord, C, Luyster, R., Gotham, K., & Guthrie, W. (2012). Autism Diagnostic Observation Schedule–Toddler Module. *Los Angeles: Western Psychological Services*.

Lord, Catherine, Elsabbagh, M., Baird, G., & Veenstra-Vanderweele, J. (2018). Autism spectrum disorder. *The Lancet*.

Luyster, R., Gotham, K., Guthrie, W., Coffing, M., Petrak, R., Pierce, K., … Oti, R. (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.

Marlow, M., Servili, C., & Tomlinson, M. (2019). A review of screening tools for the identification of autism spectrum disorders and developmental delay in infants and young children: Recommendations for use in low- and middle-income countries: Marlow et al./A review of screening tools for autism and developmental delay. *Autism Research*, *12*(2), 176–199. https://doi.org/10.1002/aur.2033

Mazurek, M. O., Handen, B. L., Wodka, E. L., Nowinski, L., Butter, E., & Engelhardt, C. R. (2014). Age at First Autism Spectrum Disorder Diagnosis: The Role of Birth Cohort, Demographic Factors, and Clinical Features. *Journal of Developmental & Behavioral Pediatrics*, *35*(9), 561–569.

McPheeters, M. L., Weitlauf, A., Vehorn, A., Taylor, C., Sathe, N. A., Krishnaswami, S., … Warren, Z. E. (2016). *Screening for Autism Spectrum Disorder in Young Children*.

Miller, J. S., Gabrielsen, T., Villalobos, M., Alleman, R., Wahmhoff, N., Carbone, P. S., & Segura, B. (2011). The each child study: Systematic screening for autism spectrum disorders in a pediatric setting. *Pediatrics*, peds. 2010-0136.

Morales, M., Mundy, P., Delgado, C. E. F., Yale, M., Messinger, D., Neal, R., & Schwartz, H. K. (2000). Responding to Joint Attention Across the 6- Through 24-Month Age Period and Early Language Acquisition. *Journal of Applied Developmental Psychology*, *21*(3),

283–298. https://doi.org/10.1016/S0193-3973(99)00040-4

Mundy, P., Sigman, M., & Kasari, C. (1994). Joint attention, developmental level, and symptom presentation in autism. *Development and Psychopathology*, *6*(3), 389–401. https://doi.org/10.1017/S0954579400006003

Nah, Y.-H., Young, R. L., & Brewer, N. (2018). Development of a brief version of the Autism Detection in Early Childhood. *Autism*, 1362361318757563.

Nilsson Jobs, E., Bölte, S., & Falck-Ytter, T. (2018). Spotting Signs of Autism in 3-Year-Olds: Comparing Information from Parents and Preschool Staff. *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s10803-018-3821-5

Øien, R. A., Schjølberg, S., Volkmar, F. R., Shic, F., Cicchetti, D. V., Nordahl-Hansen, A., … Chawarska, K. (2018). Clinical Features of Children With Autism Who Passed 18-Month Screening. *Pediatrics*, *141*(6). Retrieved from http://pediatrics.aappublications.org/content/pediatrics/141/6/e20173596.full.pdf

Ozonoff, S., Iosif, A.-M., Baguio, F., Cook, I. C., Hill, M. M., Hutman, T., … Sigman, M. (2010). A prospective study of the emergence of early behavioral signs of autism. *Journal of the American Academy of Child & Adolescent Psychiatry*, *49*(3), 256-266. e2.

Ozonoff, S., Young, G. S., Belding, A., Hill, M., Hill, A., Hutman, T., … Schwichtenberg, A. J. (2014). The broader autism phenotype in infancy: When does it emerge? *Journal of the American Academy of Child & Adolescent Psychiatry*, *53*(4), 398-407. e2.

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. https://doi.org/10.1111/jcpp.12421

Pierce, K., Gazestani, V. H., Bacon, E., Barnes, C. C., Cha, D., Nalabolu, S., … Courchesne, E. (2019). Evaluation of the Diagnostic Stability of the Early Autism Spectrum Disorder Phenotype in the General Population Starting at 12 Months. *JAMA Pediatrics*, *173*(6), 578. https://doi.org/10.1001/jamapediatrics.2019.0624

Reilly, S., Bavin, E. L., Bretherton, L., Conway, L., Eadie, P., Cini, E., … Wake, M. (2009). The Early Language in Victoria Study (ELVS): A prospective, longitudinal study of communication skills and expressive vocabulary development at 8, 12 and 24 months. *International Journal of Speech-Language Pathology*, *11*(5), 344–357. https://doi.org/10.1080/17549500903147560

Richards, M., Mossey, J., & Robins, D. L. (2016). Parents' concerns as they relate to their child's development and later diagnosis of autism spectrum disorder. *Journal of Developmental and Behavioral Pediatrics: JDBP*, *37*(7), 532.

Risi, S., Lord, C., Gotham, K., Corsello, C., Chrysler, C., Szatmari, P., … Pickles, A. (2006). Combining information from multiple sources in the diagnosis of autism spectrum disorders. *Journal of the American Academy of Child & Adolescent Psychiatry*, *45*(9), 1094–1103.

Robins, D. L., Casagrande, K., Barton, M., Chen, C.-M. A., Dumont-Mathieu, T., & Fein, D. (2014). Validation of the modified checklist for autism in toddlers, revised with follow-up (M-CHAT-R/F). *Pediatrics*, *133*(1), 37–45.

Rosenberg, R. E., Landa, R., Law, J. K., Stuart, E. A., & Law, P. A. (2011). Factors affecting age at initial autism spectrum disorder diagnosis in a national survey. *Autism Research and Treatment*, *2011*.

Rotholz, D. A., Kinsman, A. M., Lacy, K. K., & Charles, J. (2017). Improving Early

Identification and Intervention for Children at Risk for Autism Spectrum Disorder. *Pediatrics*. Retrieved from http://pediatrics.aappublications.org/content/pediatrics/early/2017/01/10/peds.2016-1061.full.pdf

Self, T. L., Coufal, K., & Parham, D. F. (2010). Allied healthcare providers' role in screening for autism spectrum disorders. *Journal of Allied Health*, *39*(3), 165–174.

Shephard, E., Bedford, R., Milosavljevic, B., Gliga, T., Jones, E. J., Pickles, A., … Baron□Cohen, S. (2018). Early developmental pathways to childhood symptoms of attention□deficit hyperactivity disorder, anxiety and autism spectrum disorder. *Journal of Child Psychology and Psychiatry*.

Shumway, S., Farmer, C., Thurm, A., Joseph, L., Black, D., & Golden, C. (2012). The ADOS calibrated severity score: Relationship to phenotypic variables and stability over time. *Autism Research*, *5*(4), 267–276.

Stenberg, N., Bresnahan, M., Gunnes, N., Hirtz, D., Hornig, M., Lie, K. K., … Reichborn□Kjennerud, T. (2014). Identifying children with autism spectrum disorder at 18 months in a general population sample. *Paediatric and Perinatal Epidemiology*, *28*(3), 255–262.

Suren, P., Bakken, I. J., Aase, H., Chin, R., Gunnes, N., Lie, K. K., … Stoltenberg, C. (2012). Autism spectrum disorder, ADHD, epilepsy, and cerebral palsy in Norwegian children. *Pediatrics*, *130*(1), e152-158. https://doi.org/10.1542/peds.2011-3217

Williams, K. (2016). Timely identification of children with autism: Are we asking the right question? *Developmental Medicine & Child Neurology*, *58*(4), 324–325.

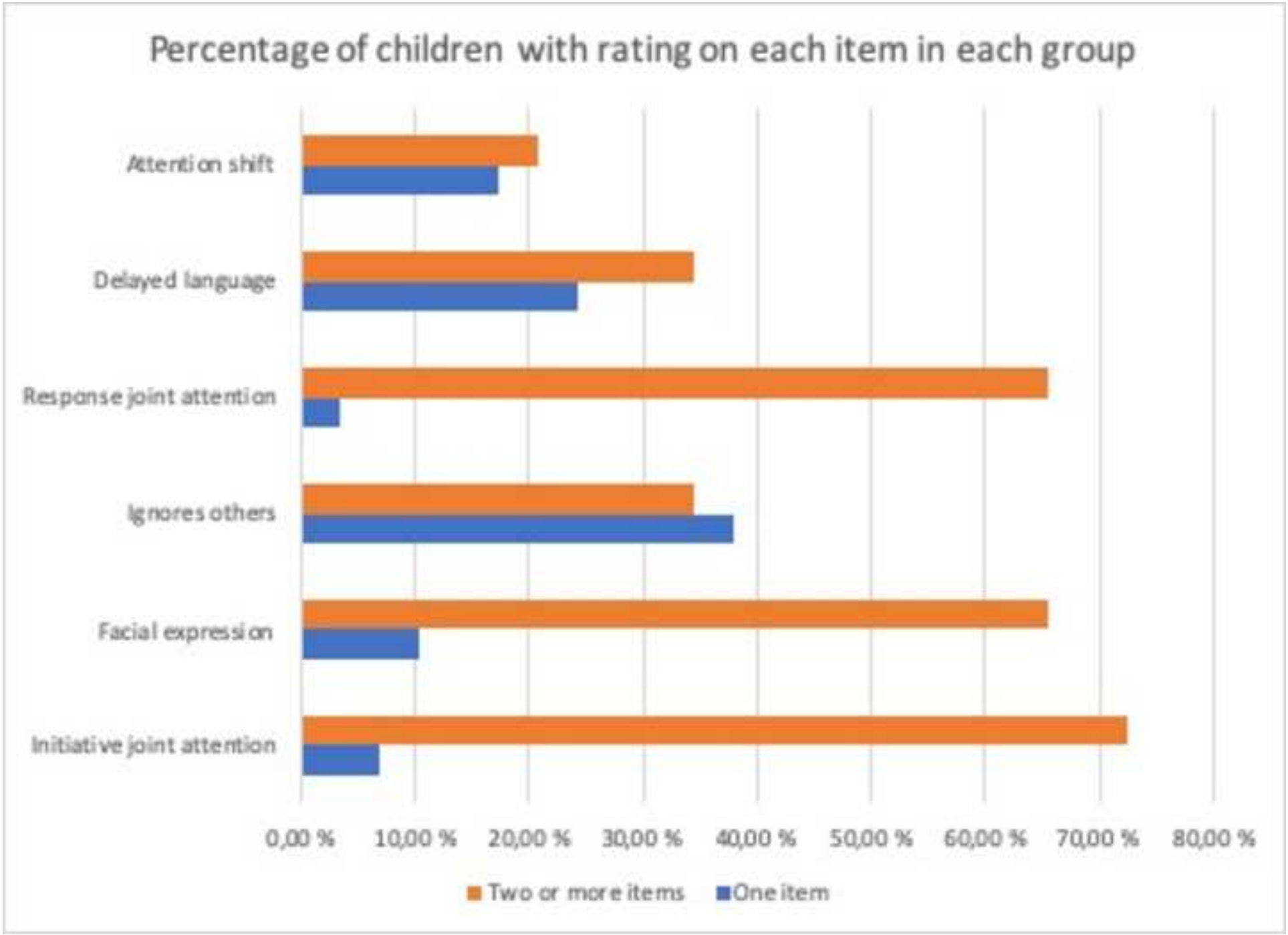
Zwaigenbaum, L., Bryson, S. E., Brian, J., Smith, I. M., Roberts, W., Szatmari, P., … Vaillancourt, T. (2016). Stability of diagnostic assessment for autism spectrum disorder between 18 and 36 months in a high-risk cohort: Early diagnostic stability in high-risk cohort. *Autism*

Research, *9*(7), 790–800. https://doi.org/10.1002/aur.1585

Zwaigenbaum, L., & Penner, M. (2018). Autism spectrum disorder: Advances in diagnosis and evaluation. *BMJ*, *361*. https://doi.org/10.1136/bmj.k1674

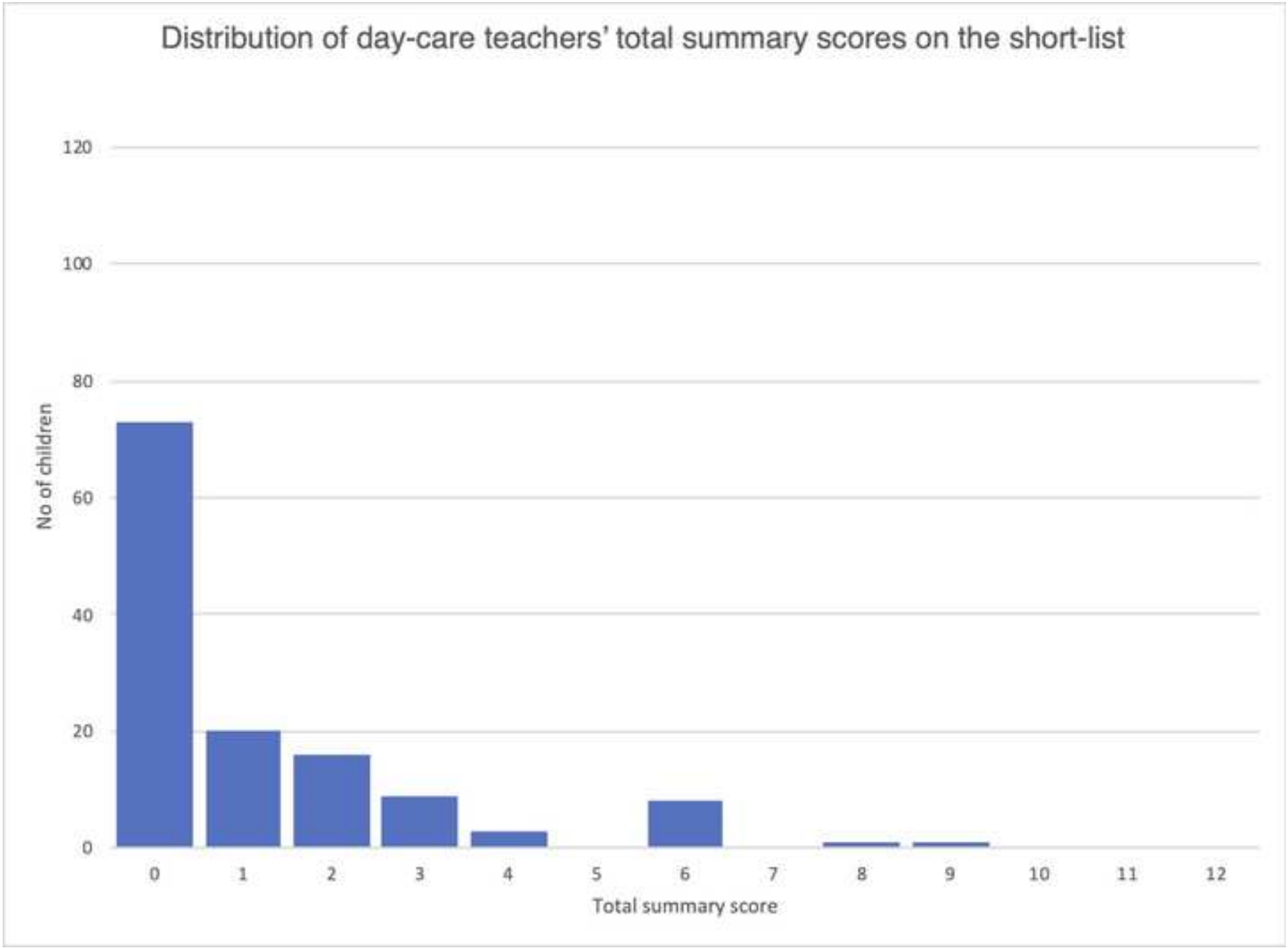
Figure

Click here to download Figure Figure 3.jpg



Figure

Click here to download Figure Figure 1 .jpg



Figure

Click here to download Figure Figure 2.pdf

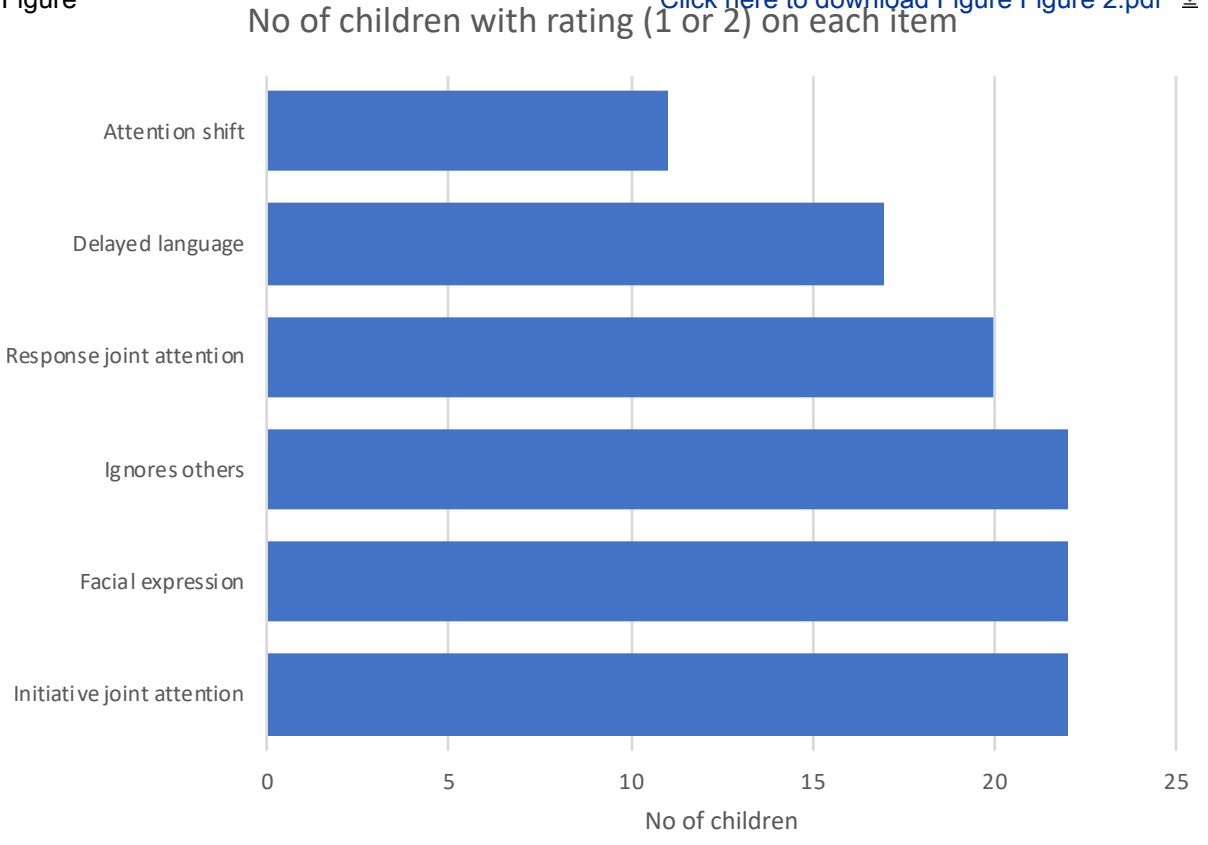


Table 1 – Demographic information on participating children.

	All participants	Participants assessed with ADOS-2	Group 1 No items	Group 2 One item	Group 3 Two or more items
	N=131	N=68 (of 89 invited)	N=24	N=20	N=24
Mean Age in months (SD)	19.66 (3.8)	19.59 (3.50)	20.1 (3.1)	19.9 (3.2)	18.8 (4.0)
Sex					
Male (%)	73 (55.7)	37 (54.4)	12 (50.0)	12 (60.0)	13 (54.2)
Female (%)	58 (44.3)	31 (45.6)	12 (50.0)	8 (40.0)	11 (45.6)
Other diagnosis (%)	7 (5.3)	4 (5.9)	3 (12.8)	1 (5.0)	0 (0.0)
Parent concerned (%)	7 (5.3)	4 (5.9)	1 (4.1)	3 (15.0)	0 (0.0)
Community health professional concerned (%)	5 (3.8)	2 (2.9)	0 (0)	1 (5)	1 (4.2)
ASD in family (%)	9 (6.9)	6 (8.8)	3 (12.5)	1 (5)	2 (8.3)
Mothers educational level					
Secondary school	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
High School	28 (21.3)	12 (17.6)	6 (25.0)	4 (20.0)	2 (8.3)
University	89 (67.9)	54 (79.4)	17 (70.8)	15 (75.0)	22 (91.7)
Missing	14 (10.7)	2 (2.9)	1 (4.2)	1 (5.0)	0 (0.0)
Fathers educational level					
Secondary school	5 (3.8)	2 (2.9)	0 (0.0)	1 (5.0)	1 (4.2)
High School	47 (35.8)	26 (38.2)	11 (45.8)	10 (50.0)	5 (20.8)
University	62 (47.3)	35 (51.5)	11 (45.8)	6 (30.0)	18 (75.0)
Missing	17 (12.9)	5 (7.4)	2 (8.3)	3 (15.0)	0 (0.0)
Income					
Low	10 (7.6)	3 (4.4)	1 (4.2)	1 (5.0)	1 (4.2)
Medium	62 (47.3)	36 (57.1)	13 (54.2)	12 (60.0)	11 (45.8)
High	41 (31.3)	24 (35.3)	7 (29.2)	5 (25.0)	12 (50.0)
Missing	18 (13.7)	5 (7.4)	3 (12.5)	2 (10.0)	0 (0.0)

Table 2. Items in the shortlist

1. Did the child develop language skills later than his or her peers?
2. When the child was playing alone, was it difficult for him or her to stop and start on another activity?
3. Did the child change facial expressions to match what was going on around him or her?
4. When you pointed at and looked at something exciting did the child look at what you pointed at and then look back at you to confirm that he or she had seen the same as you?
5. If the child saw something surprising and exciting outside the window or fence, did he or she point at it to show you?
6. When other children tried to make contact did the child often ignore them and continue with what he or she was doing?

Table 3. ADOS total score for each group of participants assessed with ADOS-2.

	All N=68	Group 1 No item N=24	Group 2 One item N=20	Group 3 Two or more items, N=24
Mean	1.82	1.16	1.20	3.00
SD	2.98	1.55	1.93	4.26
Range	0 - 19	0 - 6	0 - 8	0 - 19

Table 4. Agreement between parents and day-care professionals on individual items

	ICC (Confidence Interval)
Item 1	.497 (.334 - .623)
Item 2	.004 (-.179 - .188)
Item 3	.056 (-.13 - .238)
Item 4	-.008 (-.192 - .177)
Item 5	.155 (-.029 - .329)
Item 6	.082 (-.103 - .262)

Author Note

Dear editor,

Manuscript Number: JADD-D-19-00636

We thank you for the opportunity to review our manuscript and thank the reviewers for the useful comments on our original manuscript. We have addressed the comments from the reviewers, and below we outline our responses to the comments (*in italics*).

Reviewer #1:
I congratulate the authors on their efforts. This is an important area of inquiry. The study has several strengths, including the sample size and use of ADOS. These are interesting results. A relatively large number of children at this age display the behavior included in the short observation list. I am not familiar with the instrument's developement, but this is good information to have.

There are no fundamental flaws. I think it is fine to publish null results, but in this situation, because it is not an established tool, it would have been nice to have more data for the readers. For instance, some follow-up data on the children or revised cutoff scores for the short observation list. It would have also been nice to see more done with these data. For instance, scores by ADOS categories or characteristics of children who have high/low scores.

> *We thank the reviewer for the comments, and have included more date in the manuscript, se also responses to reviewer 2.*

Reviewer #2:
This study investigates preliminary psychometric properties of a short checklist (short-list) for observing ASD symptoms in young children (ages 12-24 months) in the daycare setting. Specifically, the authors explored: 1) the distribution of short-list scores (as rated by daycare teachers), 2) the association between short-list scores and scores on the ADOS-2, a gold standard ASD diagnostic tool, and 3) agreement of short-list ratings between daycare teachers and parents. Findings were not consistent with those previously reported using a retrospective design in a sample of children already diagnosed with ASD. In this study, the authors conclude that the short-list over identified concern for ASD symptoms, did not show agreement with ADOS-2 scores, and demonstrated poor inter-rater reliability between parents and daycare teachers.

While extending developmental screening practices to the daycare setting is a worthwhile endeavor and further research in this area could lead to important advances in early detection, several methodological problems in the current study significantly limit interpretations of the findings. First, given the reported expected prevalence of ASD in the sample (1-2 participants), one cannot evaluate performance of a measure designed to differentiate those young children who may or may not be at-risk for ASD. Second, while the ADOS-2 is used as a measure of convergent validity, there has been no judgement of clinical diagnosis from which to compare

the short-list classification to. In order to evaluate the psychometric properties of the short-list, at least with regard to how it may or may not predict risk for ASD, a group of children with confirmed ASD diagnosis is needed.

> *We acknowledge the methodical difficulties in this study. However, in this phase the objective was, as a initial exploration, to look at associations with standardized assessment. Further the inclusion of the group of children with a confirmed ASD-diagnosis was done in a previous retrospective study. We have also refocused the manuscript to clarify that the objective is not foremost to evaluate the psychometric properties.*

In addition to addressing these methodological concerns, the authors may consider the following:

Introduction:
1. A more detailed review of symptoms of ASD in 12-24 month old would be helpful to put short-list items in context.

> *We have added a more thorough description of early symptoms on early symptoms of ASD on page 2 and 3.*

2. The Introduction would benefit from a brief discussion of ASD diagnostic stability (e.g., Peirce et al., 2019, JAMA Pediatrics; Zwaigenbaum et al., 2016, Autism Research; Ozonoff et al., 2015, Journal of Child Psychology & Psychiatry) in order to further underscore the importance of early screening and detection.

> *Thank you for pointing this out, a brief discussion is added on page 3.*

3. The authors may consider reviewing the current best-practice screening instruments to more clearly describe the rationale for development of their measure.

> *We agree on the usefulness of a review of current instruments, however we consider this done by others (Marlow, Servili, & Tomlinson, 2019), and have included a discussion of this in the manuscript. Further we have discussed the challenges with current screening instruments in a Norwegian context.*

4. The authors may consider that daycare teachers often have very solid knowledge and experience with typical child development, and this may help them to be excellent and reliable observers of symptoms consistent with deviant (ASD) development.

> *Excellent point, we have discussed this now on page 5.*

5. While the authors identify their objective as "the second part in developing and exploring the feasibility of a short observation list," it seems that they are actually exploring preliminary psychometric properties of the measure.

We totally agree on this point and thank the reviewer for pointing this out. The revised manuscript have specified the objective to clarify this.

Method:
1. The authors may clarify what a "regular" daycare center is (p 5, line 55).

We have now clarified and expanded this section.

2. Why was it not ethically justified to include more children in the study? Was it not possible to assess a larger sample with ADOS-2? Further, how did the authors conclude that the short-list did not separate well between children at risk and children not at risk? Was this decision made based on analysis of the data (i.e., using ADOS-2 correlations)? This is not clear.

The decision to terminate the inclusion of new participants, was based on the rating on the short-list, and the proposed cut-off from the previous retrospective study. We have clarified this point at page 7.

3. The Toddler Module of the ADOS-2 is designed to be used with children 12-30 months with no speech to use of simple phrases. Children 12-30 months with flexible three-unit phrases should be given a Module 2 (regardless of age). It would be hard to imagine that no child included in the study had flexible phrase speech. A priori decision to only administer the Toddler Module is problematic as prior research has demonstrated that module choice affects accuracy of classification.

We recognize this challenge, as this age is characterized by rapid development in several areas, including language. A research-training ADOS professional evaluated the appropriateness of assessing each child. However, this challenge is now further discussed in the limitations section of the manuscript.

4. It would be helpful to list the full items in the short-list (in Measures).

We have added the full items as a table.

5. Table 1: how was "other diagnosis" measured?

The reporting of eventual other diagnosis was done by parents using a questionnaire developed in this study. This is now clarified in the manuscript.

Results:
1. Even in group 3, ADOS-2 scores are low. Did any participants have a classification in the range

of concern for ASD? Given the range of scores, it seems yes. For these children, albeit likely a very small n, did the short-list demonstrate elevated concern?

Only one participant scored over cut-of indicating a possible ASD on the ADOS-2. This participant also had the highest score on the short-list. However, after discussion with day-care teachers, parents and a clinical observation, no indication of ASD was found. The evaluated scores were most likely influenced by other, not developmental, factors.

2. What was the conceptual reasoning for exploring correlations between all short-list items and ADOS-2 items? Was this analysis hypothesis driven? If just exploratory, what is meaning of the significant associations (e.g., correlation between "initiation of joint attention" and "anxiety")?

The exploration of correlations between the short-list items and ADOS-2 items was hypothesis driven. We sought to explore items in the short-list that logically may be related to ADOS-2 items based on their definition in ADOS-2. The rationale of exploring anxiety was that elevated levels of anxiety may influence the frequency and quality of initiations of joint attention. The analysis is now more thoroughly described in the manuscript.

3. Short-list items are not consistently referred to in the same way (e.g., "initiates to joint attention" and "initiation to joint attention" - is this the same item?).

This is a typo, and we appreciate making us aware of this. It refers to the same item, and is now corrected in the manuscript.

4. Tables are not labeled consistently (e.g., Table 1, 2, 5).

Now corrected in the revised manuscript.

5. Figures are not labeled clearly; need figure captions to understand what is being reported.

Captions are now added.

6. Figure representing "Number of children with ratings of 1 or 2 on each item" - would be helpful to see how this differs among the groups.

We have added a new figure showing this (figure 3).

Discussion:
1. What would the short-list add above and beyond current screening instruments; for example, why not adapt the MCHAT for use in the daycare setting?

The clarify our rationale for exploring the possibility to develop a new tool for day-care centers are now expanded in the paragraph describing current challenges in screening. We emphasize the issues raised by others on using current tools on a population level, as well at the lack of knowledge on cultural factors related to translating other screening tools.

2. The authors state that "12-24 months may be too early to detect deficits …by screening a population-based sample of children" - however, there is solid evidence that the MCHAT does this well for children 16-30 months.

There exists considerable research on the use of MCHAT to detect deficits in this age period. However, recent research on population level have raised concerns that may call for continued development of screening tools (Øien et al., 2018; Stenberg et al., 2014, Guthrie et al., 2019), as now pointed out in the manuscript.

3. Page 12, line 46, why is this data not shown?

We have now included this result in the "results" in the revised manuscript.

4. While the idea that providing training to daycare workers so they may better recognize the skills assessed on the short list makes good sense, this also makes the measure less efficient and cost effective for screening large numbers of children.

We thank the reviewer for pointing this out, and we have now including this issue in the revised manuscript.

5. Page 13, line 9-12, the authors state, "inclusion of repeated time points of measurement over a longer period of time may have improved the sensitivity and specificity of the results" - sensitivity and specificity was not measured in this study.

The purpose of this study was as the reviewer point out not to measure the sensitivity and specificity, but to explore correlations with a standardized measure. We have rewritten this sentence, to be in line with the objective of the study.

6. While the challenges with implementation of ASD screening reviewed (e.g., p 14, lines 36-53) are important, discussion of this seems premature given that the study did not include a group of children who had a confirmed ASD diagnosis. There is a focus on discussion of false negatives, though it seems that concern for false positives was more relevant to findings of the current study.

We have expanded the discussion on screening, to be further discuss the issue of false positives.

7. One limitation not mentioned is the problem of self-selection bias; is there something different about those who agreed to follow-up ADOS-2 assessment than those who didn't. This cannot necessarily be overcome in a study like this; however, it is a limitation that may be worth mentioning.

> *We again thank you for raising this important issue, it is now mentioned in the discussion.*

Minor points to address:
1. There are many grammatical errors throughout the manuscript (e.g., p 2, line 31; p 8, line 19).

> *The manuscript has been checked and revised to improve the language and reduce the grammatical errors.*

We hope that you will consider our revised manuscript for publication and are positive to perform further revisions if needed.

On behalf of the authors
Yours sincerely

Kenneth Larsen