ASHA American Speech-Language-Hearing Association

JSLHR

Research Article

Assessing Social Communication and Measuring Changes in Chinese Autistic Preschoolers: A Preliminary Study Using the Social Communication Scale

Li Wang,[a,b] Xin Qi,[c] Ziyan Meng,[d] Meiyu Xiang,[b] Zhuoqing Li,[b] Sitong Zhang,[b] Longyun Hu,[b] Hoyee W. Hirai,[b] Carol K. S. To,[c] and Patrick C. M. Wong[b,d]

[a]Shenzhen Research Institute, The Chinese University of Hong Kong, Shenzhen, China [b]Brain and Mind Institute, The Chinese University of Hong Kong, Hong Kong SAR, China [c]Faculty of Education, The University of Hong Kong, Hong Kong SAR, China [d]Department of Linguistics and Modern Languages, The Chinese University of Hong Kong, Hong Kong SAR, China

ARTICLE INFO

Article History:
Received April 16, 2024
Revision received July 21, 2024
Accepted January 2, 2025

Editor-in-Chief: Julie A. Washington
Editor: Cathy Huaqing Qi

https://doi.org/10.1044/2025_JSLHR-24-00255

ABSTRACT

Purpose: Assessing social communication and measuring its changes among young autistic children presents significant challenges, particularly when tracking intervention effects within short timeframes. Existing measures, mostly validated in Western contexts, may not be suitable for culturally diverse populations. Addressing this gap, the Social Communication Scale (SCS) was developed to provide a culturally accessible and reliable measure for the Chinese population. This study explores the psychometric properties of the SCS and its ability to capture intervention-induced changes.
Method: Fifty-two autistic children aged 2–5 years were recruited from China. One parent per family participated in a 20-week support program aimed at enhancing parents' communication strategies to prompt social communication with their children at home. The SCS was administered before and after the program.
Results: The SCS exhibited outstanding overall interrater reliability (ICC = .91) and convergent validity with established measures, including the Autism Diagnostic Observation Schedule–Second Edition, the Communication subdomain of the Mullen Scales of Early Learning, and the Vineland Adaptive Behavior Scales–Third Edition. Notably, the SCS effectively captured subtle changes during the 20-week intervention.
Conclusions: As the first social communication scale developed for Chinese autistic preschoolers, the SCS proves to be a reliable and valid measure. This addresses unique challenges in autism assessment and intervention in China. To strengthen its broader applicability, future research should prioritize validating the SCS with larger and more diverse samples across various regions, contributing to a comprehensive understanding of its value and limitations.
Supplemental Material: https://doi.org/10.23641/asha.28569035

Autism spectrum disorder (ASD) is a lifelong neurodevelopmental condition that typically emerges in early childhood (American Psychiatric Association, 2013). A defining characteristic of autistic individuals is atypical social communication and interaction. For parents of autistic children, their primary concern is mitigating the impact of the condition on their children's daily lives. An accumulation of research evidence indicates that timely, evidence-based approaches play a crucial role in enhancing the social and communicative skills of autistic children. Such interventions can enhance children's developmental and learning trajectories and meet family needs such as their well-being (Fuller & Kaiser, 2020; Grzadzinski et al., 2020; Stahmer et al., 2019).

Correspondence to Patrick C. M. Wong: p.wong@cuhk.edu.hk. ***Disclosure:*** *The authors have declared that no competing financial or nonfinancial interests existed at the time of publication.*

Challenges in Evaluating Social Communication Support Programs

Evaluating the effectiveness of social communication support programs can be challenging due to the subtle, diverse, and individual nature of improvements in social communication behaviors, especially over short periods (Grzadzinski et al., 2020). Outcome measures lacking sensitivity may fail to capture these nuanced shifts. Additionally, these measures must be multidimensional, capturing both quantitative and qualitative aspects to provide a holistic understanding of progress.

Current Outcome Measures

Recognizing the necessity for robust and valid outcome measures to discern evidence-based approaches, Autism Speaks convened a panel of experts in 2011 to assess 32 existing measures for their suitability as social communication outcome measures (Anagnostou et al., 2015). While none were deemed free of limitations, six measures were identified as appropriate for specific uses with caveats, and six others were considered potentially appropriate but lacked sufficient supporting evidence at the time. In addition, Grzadzinski et al. (2020) recommended four newly developed or revised measures for quantifying intervention responses in autistic toddlers and preschoolers. In total, 16 measures were identified as either appropriate or potentially suitable for use (see Table 1). However, over half (n = 10) are questionnaires to be completed by caregivers, teachers, or individuals themselves. Such measures are prone to potential placebo effects arising from participation in support programs (Aman et al., 1985a, 1985b; Cen et al., 2017; Chang et al., 2013; Cohen, 2003; Constantino & Gruber, 2012; Goldstein & Naglieri, 2009; Gresham & Stephen, 2008; Hartman et al., 2006; Kanne et al., 2014; Kat et al., 2020; Oakland & Harrison, 2011; Reynolds, 2010; Sparrow et al., 2005; Tan et al., 2020; Z. Wu et al., 2019; Zhang et al., 2020; B. Zhou et al., 2017). Other measures primarily target adaptive behaviors or cognitive functioning skills, which are not the central focus of social communication programs, even though they may have downstream impacts on more global cognitive and adaptive skills over years (Bayley, 1993; Chang et al., 2013; Cheong et al., 2022; Eapen et al., 2013; Hua et al., 2019; Mullen, 1995; Oakland & Harrison, 2011; Rogers et al., 2012; Sparrow et al., 2005; Zhang et al., 2020). Another limitation is the narrow age range targeted by many existing measures, which restricts their utility as longitudinal outcome tools, and structured play, often used in some measures, may pose additional challenges for autistic children as it can lead to higher rates of irritability and can negatively affect children's social communication performances (Fipp-Rosenfield et al., 2023; Mundy et al., 2007; Wetherby & Prizant, 2002).

Need for a New Measure in China

It is worth noting that current measures are largely developed in Western cultural contexts and may not directly apply to other cultural contexts, such as China, for several reasons. First, research highlights significant variations in social communication behaviors across cultures (Golson et al., 2022). Behaviors such as prosody, eye gaze, gestures, and integration of gaze and language vary across cultures (Lau et al., 2022; Nayar et al., 2021, 2023). Second, parent–child interactions, which serve as a basis for some recommended outcome measures, can also differ across cultures (Bornstein & Esposito, 2020). Thus, ensuring that assessments are validated within culturally distinct populations is essential for capturing the full diversity of autism manifestations.

In addition to cultural considerations, implementing current outcome measures in China is associated with substantial practical constraints (Pang et al., 2018). These include the financial burden of training personnel to administer assessments, acquiring necessary materials, and accommodating linguistic differences among testers. Measures, such as the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2012) and the Communication and Symbolic Behavior Scales: Developmental Profile Behavior Sample (Wetherby & Prizant, 2002), recommend that professional qualifications, such as training as a speech-language therapist or clinical psychologist, are prerequisites to be trained to administer the tools (Grzadzinski et al., 2016; Wetherby & Prizant, 2002). However, China faces a severe shortage of trained professionals in these fields. With a reported 298 million children between the ages of 0 and 17 years, there are fewer than 100,000 pediatricians, fewer than 500 full-time child psychiatrists, and a national scarcity of trained speech-language therapists, although exact figures are not available in China (Law et al., 2023; J. Wu & Pan, 2019). Moreover, the procurement of standard tools including the toys or other materials required for some measures, for example, ADOS-2, Brief Observation of Social Communication Change (BOSCC; Grzadzinski et al., 2016), Mullen Scales of Early Learning (MSEL; Mullen, 1995), and Bayley Scales of Infant Development (Bayley, 1993), entails individual purchases, often necessitating approved agreements, further escalating the overall cost and diminishing accessibility for Chinese official institutes. Finally, language barriers further impede the accessibility of existing measures. Most of the measurements were developed and validated in English, and the training is also predominantly available in English. Professionals in China face difficulties in

Table 1. Summary of recommended currently available outcome measures.

Study	Measures	Domain	Type	Reporter	Age
*†Reynolds (2010) Tan et al. (2020; Chinese version)	Behavior Assessment System for Children (BASC)	Social Skills, Problem Behaviors	Questionnaire	Caregiver, teacher, self	2–25 years
*†Gresham & Stephen (2008) Wu et al. (2019; Chinese version)	Social Skills Improvement System (SSIS)	Social Skills, Problem Behaviors	Questionnaire	Caregiver, teacher, self	3–18 years
†Kanne et al. (2014)	Autism Impact Measure (AIM)	Social Communication, Restricted and Repetitive Behaviors, Peer Interaction	Questionnaire	Caregiver	2–17 years
*Constantino & Gruber (2012) Cen et al. (2017; Chinese version)	Social Responsiveness Scale (SRS)	Social Skills, Autistic Preoccupations	Questionnaire	Caregiver, teacher, self	> 30 months
*Goldstein & Naglieri (2009) Zhou et al. (2017; Chinese version)	Autism Spectrum Rating Scales (ASRS)	Social Communication, Problem Behaviors,	Questionnaire	Caregiver, teacher	2–18 years
*Cohen (2003)	Pervasive Developmental Disorder–Behavior Inventory (PDD-BI)	Social Communication, Problem Behaviors	Questionnaire	Caregiver, teacher	18 months to 8 years 5 months
*Hartman et al. (2006)	The Children's Social Behavior Questionnaire (CSBQ)	Problem Behaviors	Questionnaire	Caregiver	4–18 years
*Aman et al. (1985a, 1985b) Kat et al. (2020; Chinese version)	Aberrant Behavior Checklist (ABC)	Problem Behaviors	Questionnaire	Caregiver, teacher	6–54 years
*†Sparrow et al. (2005) Zhang et al. (2020; Chinese version)	Vineland Adaptive Behavior Scales (Vineland)	Adaptive Behaviors	Questionnaire	Caregiver, Teacher	0–90 years
*†Oakland & Harrison (2011) Chang et al. (2013; Chinese version)	Adaptive Behavior Assessment System (ABAS)	Adaptive Behaviors	Questionnaire	Caregiver, teacher, self	0–89 years
†Mullen (1995) Cheong et al. (2022; Chinese version)	Mullen Scales of Early Learning (MSEL)	Cognitive Functioning	Direct observation	Examiner	0–68 months
†Bayley (1993) Hua et al. (2019; Chinese version)	Bayley Scales of Infant Development (Bayley)	Cognitive Functioning	Direct observation	Examiner	1–42 months
*Lord et al. (2012) Chen et al. (2023; Chinese version)	Autism Diagnostic Observation Schedule (ADOS)	Social Communication, Restricted and Repetitive Behaviors	Direct observation	Examiner	> 12 months
*†Wetherby & Prizant (2002) Lin et al. (2015; Chinese version)	Communication and Symbolic Behavior Scales Developmental Profile Behavior (CSBS DP-BS)	Social Communication, Symbolic Behaviors	Direct observation	Examiner	8–24 months
†Grzadzinski et al. (2016)	Brief Observation of Social Communication Change (BOSCC)	Social Communication, Restricted and Repetitive Behaviors	Video-taped observation	Caregiver, examiner	1–10 years
*†Mundy et al. (2007)	Early Social Communication Scales (ESCS)	Social Communication	Video-taped observation	Examiner	8–30 months

*Indicates recommended by Anagnostou et al. (2015). †Indicates recommended by Grzadzinski et al. (2020).

utilizing them effectively due to the language difference. Proficient English becomes a prerequisite, presenting a notable hurdle, particularly in regions where English language skills are not widely prevalent.

All these challenges collectively impede the comprehensive and equitable provision of autism support services in China. Addressing these obstacles requires concerted efforts to alleviate financial burdens, expand professional

training programs, and develop culturally sensitive measures tailored to the unique needs of the Chinese autistic population.

The Potential of Telehealth in China

The emergence of telehealth-based outcome measures, such as the BOSCC, provides promising alternatives (Grzadzinski et al., 2016). This not only saves travel time but also enhances health care accessibility, allowing individuals to participate in assessments from the convenience of their homes. This is particularly important for China, a vast country with diverse geographical challenges and a large population. Notably, the distribution of the limited pediatricians, child psychiatrists, and speech-language therapists is highly uneven, with most of them located in major cities, such as Beijing and Shanghai, further emphasizing the necessity of the telehealth use of outcome measures (J. Wu & Pan, 2019). The smooth use of telehealth heavily relies on the internet. As of June 2023, the number of internet users in China reached 1.079 billion, an increase of 11.09 million from December 2022, with the internet penetration rate reaching as high as 76.4% (Xinhua, 2023). Aligning with the growing trend of internet users, telehealth offers a more convenient and efficient way to assess autistic children across regions, ultimately contributing to improved health care and intervention outcomes.

The Present Study

In response to these pressing challenges faced in China and the unprecedented potential of telehealth, this study developed the Social Communication Scale (SCS), an observational scale designed to capture subtle changes in social communication behaviors associated with autism in Chinese preschoolers aged between 2 and 5 years over a short intervention period. Notably, the SCS is the first scale of its kind developed for Chinese preschoolers and can be conducted remotely. This study aimed to examine the initial psychometric properties of the SCS in Chinese preschoolers, including reliability and construct validity, and to provide preliminary evidence of its utility as a measure of change in social communication behaviors over time.

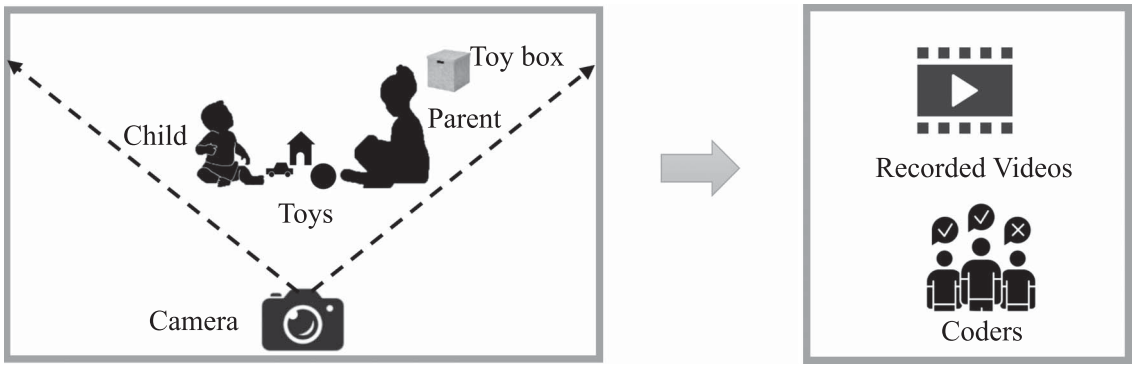
Method

Development of the SCS

The SCS is a play-based observational tool designed to assess children's social communication skills. In this scale, parents actively engage in naturalistic play activities with their children during the administration phase, while trained coders evaluate the children's social communication abilities based on recorded videos of these parent–child interactions during the coding phase, as illustrated in Figure 1.

Drawing insights from a pertinent white paper addressing telehealth evaluations of autism during the COVID-19 pandemic (Jang et al., 2022), the SCS recommends four play activities to provide different contexts for observing and assessing social communication behaviors: (a) free play for 4 min, (b) pretend play for 4 min, (c) snack for 2 min, and (d) bubble play for 2 min. It often takes approximately 15–20 min to complete, including transitions. To facilitate seamless implementation, parents are provided with explicit written instructions for utilizing the SCS. These instructions encompass (a) an introduction to the SCS and (b) critical considerations such as advanced setup of the play environment, optimal camera positioning to record the face-to-face interaction, and the preparation and storage of toys distinctly for each activity in an opaque box. In addition, the instruction lists (c) the types of toys recommended for each activity primarily

Figure 1. Illustration of Social Communication Scale procedures.



emphasizing the specific attributes of the toys (see Supplemental Material S1). For example, in the free play, pop-up toys are recommended, which can be jack-in-the-boxes or other toys with a pop-up feature, and in the bubble play, a bubble maker is suggested, which can be a bubble gun or other toys that produce bubbles. Families are not supplied with external toys for these activities. The instructions also offer (d) concise guidelines for each activity, such as parents being prompted to start by presenting the toys to the child, allowing a brief exploration period, encouraging natural play, and notifying the child when transitioning between activities. This clear and detailed guidance not only ensures consistency in evaluation but also provides a user-friendly framework for parents and coders engaged in the assessment process.

Upon completing the administration phase, coders thoroughly examine recorded videos to evaluate children's social communication skills using the SCS coding scale. This coding scale is carefully crafted in alignment with the diagnostic criteria for autism within the social communication and interaction category of the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* (American Psychiatric Association, 2013). The SCS coding scale initially delves into two subdomains (see Table 2): (A) verbal and nonverbal social interaction, considering behaviors that autistic children commonly have difficulty with, such as "pointing" and "eye contact," and (B) socioemotional reciprocity, considering behaviors that are reciprocated using individual or combinations of verbal and nonverbal means from subdomain A, such as making a request using a combination of pointing and eye contact. Each of these subdomains comprises five items, mirroring the structure of the ADOS-2 (Lord et al., 2012) and the Autism Diagnostic Interview–Revised (Rutter et al., 2003). Of significance, the scale for each item extends to a 7-point scale, and the higher the score, the more severe the social communication difficulties (see an example of one item in Supplemental Material S1). This extension is deliberate, aiming to comprehensively consider both the quality, exemplified by a definitive demonstration of the skill, and the quantity, encompassing aspects such as frequency, varieties, and applied activities. This nuanced approach ensures the sensitivity of the assessment, enabling the capture of subtle changes in the observed social communication skills. To ensure the reliability of the SCS, each item underwent refinement through multiple iterations. The goal was to eliminate potential floor or ceiling effects across different levels of social communication competence.

Table 2. The subdomains, index, and items of the Social Communication Scale.

Subdomain	Index	Item
A. Verbal and Nonverbal Social Interaction	A1	Spontaneous vocalization
	A2	Pointing
	A3	Gestures
	A4	Eye contact
	A5	Facial expressions directed to others
A. Social–Emotional Reciprocity	B1	Requesting
	B2	Social overtures
	B3	Initiation of joint attention
	B4	Social interaction
	B5	Enjoyment in interaction

Participants and Procedure

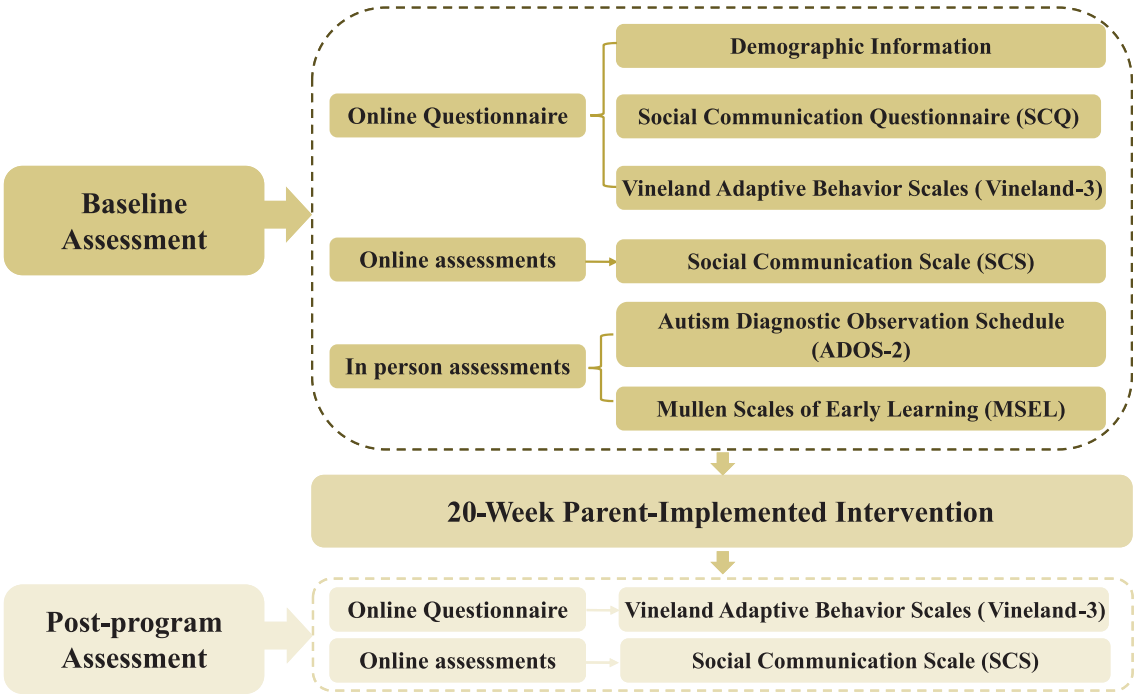
Fifty-two children aged 2–5 years who met the inclusion criteria of either a clinical diagnosis or a high likelihood of autism determined by a qualified clinician were recruited from Mainland China. Recruitment was facilitated through parent support networks, early intervention centers, referrals from pediatricians, social media platforms (e.g., WeChat), and word of mouth. These children and their parents are either Mandarin or Cantonese speakers and participated in an ongoing parent-implemented social communication support program. Children with severe hearing or visual impairment, or any neurological or psychiatric conditions requiring medications, were excluded. Parents with a history of neurological injuries, a severe psychiatric disorder requiring support, a severe hearing or visual impairment, or any condition that prohibited them from conducting the program at home were also excluded. The diagnosis of children was further confirmed using the ADOS-2 by a clinical psychologist with clinical and research reliability for administration and scoring of ADOS-2.

After confirming eligibility, families completed the baseline assessment (T1), and one parent from each family participated in the program, as illustrated in Figure 2. Upon the completion of the program, the postprogram assessment (T2) was conducted. The retention rate was as high as 90%, with five families lost to follow-up. See Table 3 for demographic characteristics of the recruited children. Informed consent was obtained from participating parents.

Parent-Implemented Social Communication Support Program

This study is part of a randomized controlled trial designed to compare the effectiveness of a parent-implemented social communication support program under two different conditions for Chinese autistic preschoolers. Specifically, the program introduces parents to seven communication strategies aimed at enhancing their children's social communication development. The strategies are (a) pause, observe, and listen; (b) imitate,

Figure 2. Flowchart of the assessment and support program process.



join in, and add; (c) make comments; (d) communication temptations; (e) take turns; (f) expansion; and (g) recast output, which are similar to those used in other established programs (Barnes et al., 1983; Bornstein et al., 2008; Carpenter, 2005; DeThorne et al., 2009; Fey et al., 1999; McGillion et al., 2017; Pepper et al., 2004). Given the cultural and linguistic differences in China, certain modifications to these strategies were necessary; however, the specifics of these modifications are beyond the scope of the present study. Parents learn these strategies in one of two conditions: either in a group study setting (eight families in each group) under the guidance of a speech and language therapist through online sessions over 20 weeks (20 classes) or in a self-study format where they receive the same learning materials without therapist guidance over the same period. In both conditions, parents are required to implement these strategies intensively during parent–child play interactions in their daily lives. In addition to the postprogram assessments (T2), participating families will be followed up at 12, 18, and 24 months; these follow-ups are currently ongoing. For the purposes of this preliminary study, a comparison of the two conditions will not be explored.

Outcome Measures

SCS

All 52 children completed the SCS at T1, and 47 of them completed it at T2. However, one video at T1 and two videos at T2 were incomplete due to missing toys (e.g., no pretend toys or bubbles), resulting in 51 T1 videos and 45 T2 videos for analysis. The SCS administration was facilitated through the Tencent Meeting & VooV Meeting application, allowing remote evaluation within the participants' homes. Before commencement, a trained facilitator (i.e., research assistants) logged into the application and went through the instructions of the SCS with the parent to ensure the proper setup of the environment, toys, face-to-face seating, and, crucially, the camera position to facilitate the smooth execution of activities. Throughout the assessment, parents engaged with their child in a natural manner, as they normally do at home. The facilitator provided timely reminders for transitioning between activities, alleviating parental concerns about timing, and recorded the administration for subsequent coding.

The coding of SCS was carried out by five independent coders (one PhD student, two master-level research assistants, and two speech-language pathologists), who remained unaware of both the program status and the specific time point. Prior to independent coding, these coders underwent intensive training on the SCS scale, utilizing a comprehensive SCS training video library. This library comprised videos featuring children exhibiting varying levels of autism spectrum–related symptoms, categorized from minimal to no evidence, low, moderate, and high

Table 3. Demographic characteristics of children at baseline.

Variable	All (*N* = 52)
Age (month)	40.29 (9.56)
Sex	
Female	8 (15%)
Male	44 (85%)
SCQ	21.19 (4.99)
ADOS-2 CSS (*N* = 49)	6.35 (1.56)
MSEL *T* score (*N* = 49)	119.41 (51.97)
Nonverbal IQ	59.51 (26.79)
Verbal IQ	59.92 (28.95)
Vineland-3 standard score	
Communication	62.14 (17.10)
Daily living skills	75.77 (11.84)
Socialization	66.14 (14.55)
Motor skills	84.40 (12.36)
Language background	
Monolingual (Mandarin)	36 (69%)
Monolingual (Cantonese)	5 (10%)
Bilingual (Mandarin and Cantonese)	6 (11%)
Bilingual (Mandarin and other Chinese dialects)	3 (6%)
Multilingual (Mandarin, Cantonese, and other Chinese dialects)	2 (4%)

Note. Continuous variables are presented as mean (standard deviation) and categorical variables as number (percentage). Other Chinese dialects include Hubei dialect, Chaozhou, Hakka, Minnan. SCQ = Social Communication Questionnaire; ADOS-2 = Autism Diagnostic Observation Schedule–Second Edition; CSS = calibrated severity score; MSEL = Mullen Scales of Early Learning; Vineland-3 = Vineland Adaptive Behavior Scales–Third Edition.

levels based on ADOS-2 calibrated severity score (CSS). Consensus codes for the training videos were established through collaboration between a psychologist (the first author, possessing clinical and research reliability in ADOS-2 and extensive autism research experience) and a clinical psychologist (the second author, with clinical and research reliability in ADOS-2 and substantial clinical practice in autism). The training encompassed a detailed review of the SCS coding scale, practical video watching and coding exercises, active participation in coding discussions with reliable coders, and receiving prompt feedback on their own coding on the training videos. It took approximately six to seven videos and around 20 hr of discussions for the coders who had no experience working with autism to achieve a consistent interrater agreement of 80% or higher across three consecutive videos with consensus codes. The agreement was assessed at the item level with a 1-point tolerance, meaning that each individual item was compared between the trainee's scores and the consensus scores, with a slight variation (within 1 point) deemed acceptable. Coders who achieved agreement of 80% or higher across three consecutive videos with consensus codes were invited to participate in the present study. To ensure robustness, experimental videos were anonymized, removing any information about program status and time point, and were randomly distributed among coders. The SCS coders were masked to the administrations and scores of other convergent validity measures.

Vineland Adaptive Behavior Scales–Third Edition

Parents of all 52 children completed the Vineland Adaptive Behavior Scales–Third Edition (Vineland-3) at T1, while T2 data were available for 41 children. The Vineland-3 evaluates adaptive behavior across communication, socialization, daily living, and motor skills domains, providing standard scores for each, as well as an overall adaptive behavior composite standard score. The Chinese version showed moderate to excellent reliability and validity (Zhang et al., 2020). The completion time for parents ranged from 30 to 60 min for this comprehensive scale.

Baseline Measures

ADOS-2

Due to COVID-19 restrictions, which was the period of the initial recruitment of this study, three families out of 52 (6%) were unable to attend in-person assessments. Thus, the ADOS-2 was administered to 49 children at T1 with the majority (98%, n = 48) receiving ADOS-2 Module 1 and one child receiving Module 2 (2%) based on their developmental and language levels. Widely acknowledged as the "gold standard" in autism assessments, the ADOS-2 is an activity-based evaluation conducted by qualified professionals over 40–60 min (Lord et al., 2012). It directly observes and assesses symptoms and behaviors associated with autism. The ADOS-2 total scores encompass raw scores for Social Affect and Restricted and Repetitive Behavior, which are subsequently transformable into a CSS ranging from 1 to 10. A score of 10 signifies the highest severity of autism spectrum–related symptoms (Duda et al., 2014; Gotham et al., 2009). Specifically, the ADOS-2 CSS corresponds to four distinct levels: 1–2 indicating minimal-to-no evidence, 3–4 denoting low severity, 5–7 representing moderate severity, and 8–10 reflecting high severity. The Chinese version showed high diagnostic specificity (.71 and 1.00) and sensitivity (.83 and .96; Chen et al., 2023).

MSEL

The MSEL was administered to 49 children at T1. It is a standardized assessment gauging learning abilities across expressive language, receptive language, visual reception, and gross and fine motor skills. An early learning composite score is derived from the fine motor, visual reception, receptive language, and expressive language

scales. Each subscale is standardized to calculate a standard *T* score, with the sum of fine motor and visual reception representing nonverbal IQ and of receptive and expressive language representing verbal IQ. The Chinese cultural adaptation and validation of MSEL indicated a good interrater reliability, internal consistency, and convergent validity (Cheong et al., 2022). The administration took about 30–60 min to complete.

Social Communication Questionnaire

All parents (n = 52) rated the Social Communication Questionnaire (SCQ) online, a 40-item, yes-or-no questionnaire designed to assess communication skills and social functioning in children potentially experiencing autism. The questionnaire is available in two forms: lifetime, examining the child's entire developmental history, and current, focusing on the last 3 months. The present study utilized the current form, which parents could complete in less than 10 min. The Chinese version showed good internal consistency, moderate test–retest reliability, and high diagnostic sensitivity (.85–.96) and specificity (.88–.98; Liu et al., 2022). A total score of 11 (for children under 4years old) or 15 (children 4years old and above) or above indicates a possible autism.

Data Analysis

All statistical analyses were conducted using RStudio (R Version 4.4.1; R Core Team, 2024).

Item Analysis

Item analysis was performed using the *sjPlot* package (Daniel, 2024) to examine item difficulty and item discrimination across all 96 videos.

Reliability

To assess the internal consistency of the items, Cronbach's alpha was calculated based on the full data set of 96 videos using the *psych* package (William, 2024). For the estimation of interrater reliability, all 96 videos were independently double-coded. Intraclass correlation coefficient (ICC) was measured for each item and the total score using two-way random effects models, with higher values indicating greater reliability, using the *irr* package (Gamer et al., 2019). In addition, the agreement percentage for each item was calculated with a 1-point tolerance accepted.

Validity

Before conducting factor analysis, a Pearson correlation matrix was computed using the *rstatix* package (Kassambara, 2023) to detect potential issues of collinearity between items, based on the full data set. Two items displaying high interitem correlations (i.e., correlation coefficient > .7) were identified, and one item from each pair was subsequently removed to enhance the independence of the items (Boslaugh, 2012). The overall suitability of the data set for factor analysis was assessed using the Kaiser–Meyer–Olkin (KMO) measure and Bartlett's test of sphericity, both implemented through the *psych* package. A satisfactory KMO score (≥ .5) and a significant p value from Bartlett's test (p < .05) supported the data set's adequacy for factor analysis (Kaiser, 1974). Exploratory factor analysis (EFA) was then performed on the remaining items using the *minres* method with *promax* rotation. One- and two-factor solutions were tested, with a stable factor loading set at > 0.4 (Gorsuch, 2014; Maskey et al., 2018). Convergent validity of the SCS was assessed by calculating Pearson correlations between the SCS total score and scores from established measures, including the ADOS-2 CSS, SCQ raw scores, and MSEL cognitive *T* scores. The subdomains of established measures relating to social communication were also examined, including the Social Affect raw scores of the ADOS-2, Verbal IQ scores of MSEL, and Communication standard scores of the Vineland-3. Divergent validity was evaluated by examining correlations with unrelated constructs, such as Vineland-3 Gross and Fine Motor Skills.

Measuring Intervention Responsiveness

To investigate the SCS as an outcome measure for social communication changes over time, linear mixed-effects (LME) models were fitted to a subsample (n = 45) that completed the SCS at both T1 and T2. The LME was conducted using the *lme4* package (Bates et al., 2023), with the total score of SCS as the dependent variable, timepoint (effect-coded: T1 vs. T2) as fixed effects, and participant as the random factor. Additionally, Vineland-3 communication standard scores were measured as another outcome following the above-mentioned LME structure. Forty-one parents completed the Vineland-3 at both T1 and T2. The statistical significance of the fixed effects was estimated using the *lmerTest* package (Kuznetsova et al., 2019), which provided p values for the corresponding t tests.

Results

Item Analysis

The item analysis revealed the difficulty and discrimination indices for each item (see Table 4). Item difficulty refers to how challenging an item is for respondents. It is measured on a scale from 0 to 1, where higher values indicate easier items. Specifically, an item was considered "easy" if the index was above .90, "moderate" if it was between .20 and .90, and "hard" if it was below .20 (Sabri, 2013). All 10 items were classified as moderate.

Table 4. Psychometric properties of Social Communication Scale with 10 items.

Item	Item analysis		Internal consistency	Interrater reliability	
	Difficulty	Discrimination	Cronbach's alpha	ICC [95% CI]	Agreement percentage
A1. Spontaneous Vocalization	0.53	0.47	.81	.84 [.77, .89]	79.2%
A2. Pointing	0.69	0.47	.81	.72 [.60, .80]	74.0%
A3. Gestures	0.69	0.33	.83	.80 [.71, .86]	81.2%
A4. Eye Contact	0.44	0.59	.80	.62 [.48, .73]	66.7%
A5. Facial Expressions Directed to Others	0.47	0.54	.81	.83 [.74, .89]	79.2%
B1. Requesting	0.43	0.40	.82	.63 [.50, .74]	68.8%
B2. Social Overtures	0.47	0.68	.79	.75 [.62, .83]	70.8%
B3. Initiation of Joint Attention	0.86	0.66	.80	.52 [.36, .66]	72.9%
B4. Social Interaction	0.53	0.59	.80	.65 [.52, .75]	74.0%
B5. Enjoyment in Interaction	0.42	0.53	.81	.86 [.78, .91]	81.2%

Note. ICC = intraclass correlation coefficient; CI = confidence interval.

Item discrimination indicates how well an item differentiates between respondents with varying levels of the measured trait. Higher discrimination values suggest that an item effectively distinguishes between high- and low-performing individuals. Discrimination was classified as "very good" if the index was above .40, "reasonably good" if it was between .30 and .39, "marginal" if it was between .20 and .29 (indicating items that usually need improvement), and "poor" if it was below .19 (Sabri, 2013). Except for the "Gestures" item, which was classified as reasonably good, the rest of the items were classified as very good.

Reliability

Internal Consistency

Cronbach's alpha was calculated to assess the internal consistency of the items. The overall Cronbach's alpha was .83, indicating high internal consistency. All items demonstrated high internal consistency (all $\geq$ .79), as shown in Table 4.

Interrater Reliability

ICC and agreement percentages were measured from the 96 double-coded videos. The overall interrater reliability for the total score demonstrated excellent agreement, with an ICC value of .91 (95% confidence interval [CI] [.87, .94]). An ICC value above .90 typically indicates excellent reliability, ICC values of .75–.90 indicate good reliability, ICC values of .50–.75 indicate moderate reliability, and an ICC value below .50 indicates poor reliability (Koo & Li, 2016). At the individual item level, as shown in Table 4, the ICC values ranged from .52 to .86. Agreement percentages that indicate the extent to which raters give the scores within 1-point difference to the same item ranged from 66.7% to 81.2%. These values reflect varying degrees of agreement between the raters.

Validity

Collinearity

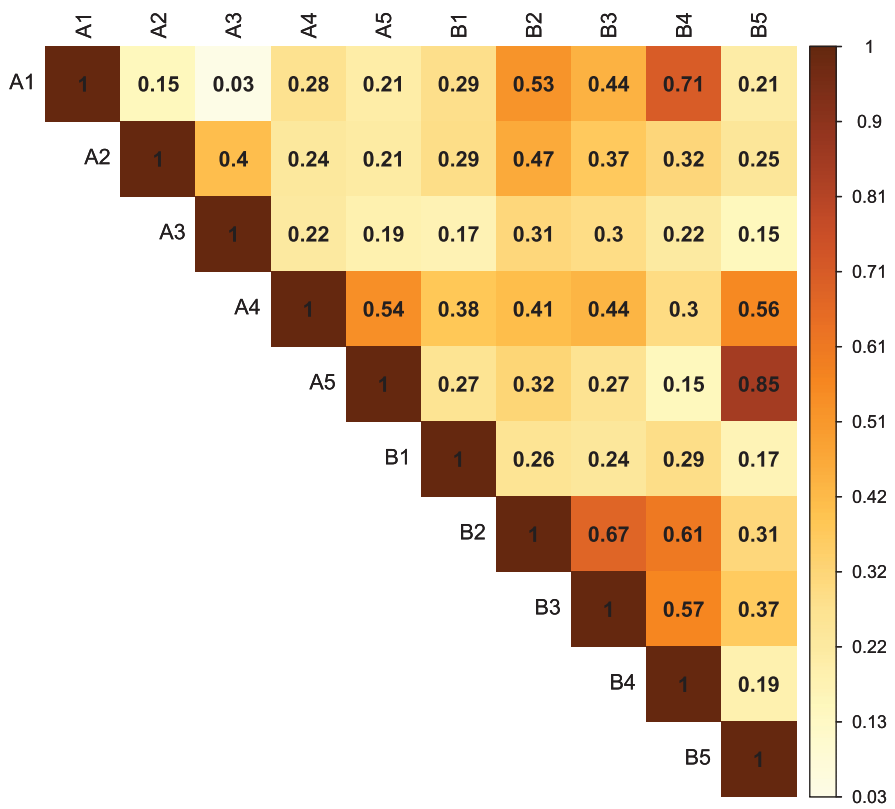
As depicted in Figure 3, the correlation between A1. Spontaneous Vocalization and B4. Social Interaction, as well as between A5. Facial Expressions Directed to Others and B5. Enjoyment in Interaction, exceeded .7, indicating a substantial overlap. To reduce redundancy and the number of items, we decided to retain the items with relatively higher interrater reliability (as indicated by higher ICC values). Specifically, since B4 and A5 had lower ICC values compared to A1 and B5, we choose to retain Spontaneous Vocalization and Enjoyment in Interaction, while removing Facial Expressions Directed and Social Interaction from the scale and further analysis.

Factor Analysis

The KMO measure was .76 overall, ranging from .72 to .83 across the remaining eight items, demonstrating a satisfactory level of adequacy for factor analysis. The Bartlett's test of sphericity, $\chi^2(28) = 219.19$, $p < .001$, against the null hypothesis of an identity matrix, supported the presence of underlying factors.

In the EFA, using the minres method with promax rotation, a one-factor solution was identified, indicating a significant single factor with standardized loadings. The SS loadings were found to be 2.72, and the proportion variance accounted for was .34. The model fit was assessed with a root-mean-square error of approximation (RMSEA) of 0.127 (90% CI [0.09, 0.17]), which indicated a borderline satisfactory fit. Most items had factor loadings greater than

Figure 3. Correlation matrix between the Social Communication Scale items.



0.4, except for the item "Gestures" (see Table 5). Subsequently, a two-factor solution was explored, revealing two factors with a correlation of .65. While the model fit was satisfactory (RMSEA = 0.096, 90% CI [0.03,0.16]), only two items were grouped into Factor 2 and the initial A and B subdomains were mixed. Therefore, we retain the one-factor solution. Due to the low factor loadings of "Gestures," the item was removed from the scale, which resulted in a slight improvement in the one-factor solution (SS loadings = 2.58, 37% of the variance, RMSEA = 0.14).

Convergent Validity

To evaluate the convergent validity of the SCS with established measures focusing on social communication, the correlation between the SCS total score and the ADOS-2 CSS, ADOS-2 Social Affect subdomain raw score, and SCQ raw scores were analyzed. The SCS total score demonstrated significant correlations with the ADOS-2 CSS score, $r(46) = .35$, $p = .01$, and the ADOS-2 Social Affect subdomain raw score, $r(46) = .55$, $p < .001$. However, no significant associations were found between the SCS and

Table 5. One- and two-factor model factor loadings for the Social Communication Scale items.

	One-factor model	Two-factor model (promax)	
Remaining items	**Factor 1**	**Factor 1**	**Factor 2**
Spontaneous vocalization	**0.51**	**0.49**	0.02
Pointing	**0.52**	**0.55**	–0.02
Gestures	0.38	0.37	0.02
Eye contact	**0.62**	–0.19	**1.12**
Requesting	**0.42**	0.22	0.24
Social overtures	**0.80**	**0.98**	–0.13
Initiation of joint attention	**0.76**	**0.72**	0.06
Enjoyment in interaction	**0.50**	0.08	**0.52**

Note. All factor loadings > 0.4 are shown in bold.

the SCQ, $r(49) = .09$, $p = .53$. In both the ADOS-2, SCS and SCQ, higher scores indicate more severe symptoms. In addition to social communication measures, significant negative correlations were observed between the SCS total score and broader communication-related tools, including the MSEL cognitive T scores, $r(46) = -.48$, $p < .001$; the scores of verbal IQ, $r(46) = -.58$, $p < .001$; and the Vineland-3 Communication standard score, $r(49) = -.54$, $p < .001$. Higher scores on the MSEL and Vineland-3 denote better abilities, indicating that better verbal IQ and communication skills were associated with lower SCS scores, reflecting less severe symptoms.

Divergent Validity

No significant associations were found between the SCS total score and Vineland-3 gross motor skills, $r(49) = -.06$, $p = .68$; Vineland-3 fine motor skills, $r(49) = -.06$, $p = .66$; maternal age, $r(49) = -.09$, $p = .53$; or paternal age, $r(49) = -.11$, $p = .45$. These results suggest that the SCS total score is not influenced by these variables, indicating good divergent validity.

Measuring Intervention Responsiveness

The results of the SCS LME model indicated a significant main effect of timepoint, $t(44) = -2.79$, $p = .008$. Children demonstrated improved performance on the SCS at T2 ($M = 21.56$, $SD = 8.28$) compared to T1 ($M = 24.6$, $SD = 7.77$). The average rate of change in the SCS total score over the 20-week support program was −1.52, indicating a decrease in scores over the 20-week support program.

For the Vineland-3 communication change, the LME results also showed a significant main effect of timepoint, $t(40) = 3.44$, $p = .001$. Children's communication improved from T1 ($M = 61.98$, $SD = 17.47$) to T2 ($M = 69.63$, $SD = 19.73$), with an average change rate of 3.83 over time. The correlation between the changes of the SCS and Vineland-3 communication was $r(37) = -.28$, $p = .09$.

Discussion

The present study highlighted the need for a culturally sensitive and accessible measure for assessing social communication behaviors over time in Chinese autistic preschoolers. The SCS was developed, and its construction was grounded explicitly in the acknowledgment of cultural variations and the challenges associated with existing Western-developed measures. In this study, we reported the initial psychometric properties of the SCS in Chinese preschoolers and provided preliminary evidence of its utility as a measure of change. The results indicated that, as the very first social communication scale developed specifically for Chinese preschoolers, the SCS was a promising and viable outcome measure.

Analysis of the psychometric properties indicated excellent interrater reliability for the SCS, with an ICC value as high as .91. This robust interrater reliability underscored the credibility of the SCS as a consistent observational tool for measuring social communication behaviors (Gisev et al., 2013). However, given the varying degrees of item-level reliability, the item-level scores should not be reported individually until greater reliability is achieved. A noteworthy aspect of this study is the successful training of doctoral students and research assistants with no prior autism experience to effectively code for the SCS. This breakthrough paves the way for a scalable and accessible approach to assess social communication behaviors in autistic children. As highlighted in the introduction, the shortage of professionals in the autism field poses a considerable challenge to implementing evidence-based assessments and interventions in China (Leng et al., 2023; Pang et al., 2018; R. Zhou et al., 2022). While training more professional personnel is an ideal and long-term goal to cater for the pressing needs of rehabilitation services in China, training nonspecialized personnel to effectively use a validated and reliable assessment tool offers a more practical and immediate solution. It not only maximizes the use of available human resources but also provides a cost-effective means of expanding the reach of the SCS (Deng & Rattadilok, 2022; Grzadzinski et al., 2016; Naveed et al., 2019).

In demonstrating the SCS's convergent validity, total scores were compared with other established measures. When compared with the ADOS-2, a recognized "gold-standard" diagnostic assessment, the SCS exhibited significant associations with both the CSS score and the Social Affect subdomain score. Importantly, the correlation was stronger with the Social Affect subdomain score than with the CSS. This difference is likely because the CSS reflects both Social Affect and Restricted and Repetitive Behaviors—an area the SCS does not assess. In contrast, the ADOS-2 Social Affect score aligns more directly with the constructs measured by the SCS. These findings emphasize the SCS's effectiveness in capturing autism-related social communication behaviors during parent–child interactions on the expanded 7-point scale. Unlike the ADOS-2, which necessitates specific toys, the SCS recommends toy types without specifying, providing flexibility and cultural appropriateness. This approach reduces costs associated with acquiring and maintaining specific toys, making assessments more accessible for families in various locations, especially in large countries like China.

Additionally, our study found significant correlations between the SCS total score and the MSEL Verbal IQ, an examiner-administered cognitive measure, as well as the Vineland-3 communication skills, a caregiver-reported questionnaire. These results provide further evidence supporting the convergent validity of the SCS. Notably, the absence of significant associations between the SCS total score and variables such as gross motor skills, maternal age, and paternal age indicates good divergent validity. However, the SCS did not exhibit associations with the SCQ, a caregiver-reported questionnaire focusing on social communication. The finding aligns with previous studies suggesting that parent ratings alone may not sufficiently capture the severity of social communication difficulties in autistic children (Corsello et al., 2007; Guttentag et al., 2022; Lee et al., 2023). This lack of association may be attributed, in part, to the relatively younger age of the current participants (i.e., 2–5 years old). Previous research has shown that sensitivity and specificity of the SCQ tend to be lower in younger age groups (e.g., < 5 and 5–7 years old) compared to older age groups (e.g., 8–10 and > 11 years old; Corsello et al., 2007). Additionally, the lack of association may be influenced by parental understanding of autism and their level of concern. It has been suggested that reporters' knowledge and concerns can influence how they respond to self-report questionnaires (Christopher et al., 2021; Havdahl et al., 2017). Our support program aims to provide timely assistance for young children who have just been diagnosed or evaluated with an elevated likelihood for autism. As these parents are relatively new to autism, their understanding and concerns may have influenced the current SCQ results, making them less reliable indicators of the child's abilities.

The SCS also captures meaningful shifts in social communication behaviors, evidenced by a significant decrease in the total score over the 20-week intervention period (average rate of change: −1.52). This negative rate of change suggests a positive trajectory, indicating a notable alleviation in the severity of social communication symptoms. For intervention studies, participant retention is a common challenge for follow-up assessments, with relocation being a prevalent factor (Bower et al., 2014; Hill et al., 2016). The SCS exhibits flexibility in remote administration, enhancing accessibility for participants facing logistical constraints and streamlining long-term follow-up assessments. However, it is crucial to consider factors influencing the validity of remote assessments, for example, environment, internet connection, and camera angles (Gibbs et al., 2021). Hence, the SCS incorporates online checks conducted by a facilitator (i.e., research assistant) prior to assessments to ensure the maintenance of validity throughout the process.

Limitations and Future Directions

Despite the promising results, it is crucial to acknowledge certain limitations in the current study. The relatively small sample size, while yielding promising results, necessitates caution in generalizing findings, especially for the results of the factor analysis. Conducting a larger scale study would bolster the robustness of the results, providing a more comprehensive understanding of the SCS's psychometrics. Additionally, this study followed families only after the support program concluded, spanning a 20-week duration. While significant changes were observed during this period, the absence of long-term follow-up data impedes a thorough exploration of the use of the SCS as an outcome measure over an extended duration to capture long-term changes. Despite the overall excellent reliability, the observed variability at the individual item level signals areas that demand attention and improvement. Focusing on items with lower agreement will be crucial to refining and strengthening the SCS, ensuring a more consistent and reliable assessment across all aspects of social communication behaviors. As aforementioned, until good reliability is achieved for all items, the item-level scores should not be reported individually. Furthermore, while the present study focused on autistic preschoolers, future studies should employ the SCS with typically developing preschoolers and children with neurodevelopment conditions other than autism. This will not only illustrate between-groups differences in social communication development but also establish a cutoff for identifying children with social communication difficulties, enhancing its clinical utility. Finally, the time required for training presents a potential barrier to the widespread adoption of the SCS. Although it is notable that individuals with no prior training in autism were able to master the measure through 20 hr of discussion-based training, this duration may deter broader implementation. Future studies could explore strategies to streamline the training process. One such approach could involve developing online, self-paced training modules with integrated automated feedback. This would allow trainees to progress at their own pace, receive real-time guidance, and alleviate scheduling constraints, potentially reducing the overall training time without compromising quality.

Conclusions

The SCS emerges as a reliable and valid tool for assessing social communication behaviors in Chinese autistic preschoolers and can capture changes following program implementation. This study addresses the unique challenges in autism assessment and intervention in China by offering a culturally sensitive, accessible, and cost-

effective measure. The development and validation of the SCS also bridge the research gap in China, providing clinicians and researchers with a localized tool. Our ongoing work will continue to provide larger samples across multiple locations to contribute to our continued understanding of the value and limitations of the SCS.

Ethics Statement

The study was approved by the Chinese University of Hong Kong Research Ethics Committee (CREC Ref. No. 2021.523-T).

Data Availability Statement

The data that support this article will be made available from the corresponding author on reasonable request.

Acknowledgments

Our research is supported by the Research Grants Council (Hong Kong Special Administrative Region, C4024-21G, to Patrick. C. M. Wong). Additional support comes from Shenzhen Natural Science Foundation Grant (JCYJ20220531103803009, to Li Wang). We thank all the participating families; the speech-language pathologists Clara C. Y. Cheng and Sandy F. Y. Yim, who delivered the support program; and the speech-language pathologists Vicky H. C. O and Kelly W. Y. Chan, who participated in the scoring.

References

Aman, M. G., Singh, N. N., Stewart, A. W., & Field, C. J. (1985a). Psychometric characteristics of the aberrant behavior checklist. *American Journal of Mental Deficiency, 89*(5), 492–502.

Aman, M. G., Singh, N. N., Stewart, A. W., & Field, C. J. (1985b). The aberrant behavior checklist: A behavior rating scale for the assessment of treatment effects. *American Journal of Mental Deficiency, 89*(5), 485–491.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5* (5th ed.). American Psychiatric Publishing.

Anagnostou, E., Jones, N., Huerta, M., Halladay, A. K., Wang, P., Scahill, L., Horrigan, J. P., Kasari, C., Lord, C., Choi, D., Sullivan, K., & Dawson, G. (2015). Measuring social communication behaviors as a treatment endpoint in individuals with autism spectrum disorder. *Autism, 19*(5), 622–636. https://doi.org/10.1177/1362361314542955

Barnes, S., Gutfreund, M., Satterly, D., & Wells, G. (1983). Characteristics of adult speech which predict children's language development. *Journal of Child Language, 10*(1), 65–84. https://doi.org/10.1017/S0305000900005146

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., Dai, B., Scheipl, F., Grothendieck, G., Green, P., Fox, J., Bauer, A., Krivitsky, P. N., Tanaka, E., & Jagan, M. (2023). *Linear mixed-effects models using 'eigen'* (Version 1.1-35.1) [Computer software]. https://cran.r-project.org/web/packages/lme4/index.html

Bayley, N. (1993). *Bayley Scales of Infant Development: Manual*. The Psychological Corporation.

Bornstein, M. H., & Esposito, G. (2020). Cross-cultural perspectives on parent–infant interactions. In J. J. Lockman & C. S. Tamis-LeMonda (Eds.), *The Cambridge handbook of infant development: Brain, behavior, and cultural context* (pp. 805–831). Cambridge University Press. https://doi.org/10.1017/9781108351959.029

Bornstein, M. H., Tamis-LeMonda, C. S., Hahn, C.-S., & Haynes, O. M. (2008). Maternal responsiveness to young children at three ages: Longitudinal analysis of a multidimensional, modular, and specific parenting construct. *Developmental Psychology, 44*(3), 867–874. https://doi.org/10.1037/0012-1649.44.3.867

Boslaugh, S. (2012). *Statistics in a nutshell: A desktop quick reference*. O'Reilly Media.

Bower, P., Brueton, V., Gamble, C., Treweek, S., Smith, C. T., Young, B., & Williamson, P. (2014). Interventions to improve recruitment and retention in clinical trials: A survey and workshop to assess current practice and future priorities. *Trials, 15*(1), Article 399. https://doi.org/10.1186/1745-6215-15-399

Carpenter, B. (2005). Early childhood intervention: Possibilities and prospects for professionals, families and children. *British Journal of Special Education, 32*(4), 176–183. https://doi.org/10.1111/j.1467-8578.2005.00394.x

Cen, C.-Q., Liang, Y.-Y., Chen, Q.-R., Chen, K.-Y., Deng, H.-Z., Chen, B.-Y., & Zou, X.-B. (2017). Investigating the validation of the Chinese Mandarin version of the Social Responsiveness Scale in a Mainland China child population. *BMC Psychiatry, 17*(1), Article 51. https://doi.org/10.1186/s12888-016-1185-y

Chang, C.-L., Lung, F.-W., Yen, C.-F., & Yang, P. (2013). Adaptive behaviors in high-functioning taiwanese children with autism spectrum disorders: An investigation of the mediating roles of symptom severity and cognitive ability. *Journal of Autism and Developmental Disorders, 43*(6), 1347–1355. https://doi.org/10.1007/s10803-012-1684-8

Chen, M.-H., Huang, C.-F., Lin, Y.-S., Chiu, Y.-N., Gau, S. S.-F., & Wu, Y.-Y. (2023). Validation of the Mandarin Chinese version of the Autism Diagnostic Observation Schedule-2 for autism spectrum disorder. *Research in Autism Spectrum Disorders, 105,* Article 102184. https://doi.org/10.1016/j.rasd.2023.102184

Cheong, P.-L., Tsai, J.-M., Wu, Y.-T., Lu, L., Chiu, Y.-L., Shen, Y.-T., Li, Y.-J., Tsao, C.-H., Wang, Y.-C., Chang, F.-M., Huang, Y.-H., & Sun, C.-W. (2022). Cultural adaptation and validation of Mullen Scales of Early Learning in Taiwanese children with autism spectrum disorder, global developmental delay, and typically developing children. *Research in Developmental Disabilities, 122,* Article 104158. https://doi.org/10.1016/j.ridd.2021.104158

Christopher, K., Bishop, S., Carpenter, L. A., Warren, Z., & Kanne, S. (2021). The implications of parent-reported emotional and behavioral problems on the modified checklist for autism in toddlers. *Journal of Autism and Developmental Disorders, 51*(3), 884–891. https://doi.org/10.1007/s10803-020-04469-5

Cohen, I. L. (2003). Criterion-related validity of the PDD behavior inventory. *Journal of Autism and Developmental Disorders, 33*(1), 47–53. https://doi.org/10.1023/A:1022278420716

Constantino, J. N., & Gruber, C. P. (2012). *Social Responsiveness Scale–Second Edition (SRS-2): Manual*. Western Psychological Services.

Corsello, C., Hus, V., Pickles, A., Risi, S., Cook, E. H., Jr., 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*(9), 932–940. https://doi.org/10.1111/j.1469-7610.2007.01762.x

Daniel, L. (2024). *sjPlot: Data visualization for statistics in social science* [Computer software]. https://CRAN.R-project.org/package=sjPlot

Deng, L., & Rattadilok, P. (2022). The need for and barriers to using assistive technologies among individuals with autism spectrum disorders in China. *Assistive Technology, 34*(2), 242–253. https://doi.org/10.1080/10400435.2020.1757787

DeThorne, L. S., Johnson, C. J., Walder, L., & Mahurin-Smith, J. (2009). When "Simon says" doesn't work: Alternatives to imitation for facilitating early speech development. *American Journal of Speech-Language Pathology, 18*(2), 133–145. https://doi.org/10.1044/1058-0360(2008/07-0090)

Duda, M., Kosmicki, J. A., & Wall, D. P. (2014). Testing the accuracy of an observation-based classifier for rapid detection of autism risk. *Psychiatry, 4*(8), Article e424. https://doi.org/10.1038/tp.2014.65

Eapen, V., Črnčec, R., & Walter, A. (2013). Clinical outcomes of an early intervention program for preschool children with autism spectrum disorder in a community group setting. *BMC Pediatrics, 13*(1), Article 3. https://doi.org/10.1186/1471-2431-13-3

Fey, M. E., Krulik, T. E., Loeb, D. F., & Proctor-Williams, K. (1999). Sentence recast use by parents of children with typical language and children with specific language impairment. *American Journal of Speech-Language Pathology, 8*(3), 273–286. https://doi.org/10.1044/1058-0360.0803.273

Fipp-Rosenfield, H. L., Levy, R. S., Grauzer, J. M., Kaat, A., & Roberts, M. Y. (2023). Autistic children's irritability during social communication assessments. *American Journal of Speech-Language Pathology, 33*(1), 369–377. https://doi.org/10.1044/2023_AJSLP-23-00216

Fuller, E. A., & Kaiser, A. P. (2020). The effects of early intervention on social communication outcomes for children with autism spectrum disorder: A meta-analysis. *Journal of Autism and Developmental Disorders, 50*(5), 1683–1700. https://doi.org/10.1007/s10803-019-03927-z

Gamer, M., Lemon, J., & Singh, I. F. P. (2019). *irr: Various coefficients of interrater reliability and agreement* (Version 0.84.1) [Computer software]. https://cran.r-project.org/web/packages/irr/index.html

Gibbs, V., Cai, R. Y., Aldridge, F., & Wong, M. (2021). Autism assessment via telehealth during the COVID-19 pandemic: Experiences and perspectives of autistic adults, parents/carers and clinicians. *Research in Autism Spectrum Disorders, 88,* Article 101859. https://doi.org/10.1016/j.rasd.2021.101859

Gisev, N., Bell, J. S., & Chen, T. F. (2013). Interrater agreement and interrater reliability: Key concepts, approaches, and applications. *Research in Social and Administrative Pharmacy, 9*(3), 330–338. https://doi.org/10.1016/j.sapharm.2012.04.004

Goldstein, S., & Naglieri, J. A. (2009). *Autism Spectrum Rating Scales (ASRS)*. Multi-Health System.

Golson, M. E., Ficklin, E., Haverkamp, C. R., McClain, M. B., & Harris, B. (2022). Cultural differences in social communication and interaction: A gap in autism research. *Autism Research, 15*(2), 208–214. https://doi.org/10.1002/aur.2657

Gorsuch, R. L. (2014). *Factor analysis: Classic edition* (2nd ed.). Routledge. https://doi.org/10.4324/9781315735740

Gotham, K., Pickles, A., & Lord, C. (2009). Standardizing ADOS scores for a measure of severity in autism spectrum disorders. *Journal of Autism and Developmental Disorders, 39*(5), 693–705. https://doi.org/10.1007/s10803-008-0674-3

Gresham, F., & Stephen, N. E. (2008). *Social Skills Improvement System (SSIS) Rating Scales*. Pearson.

Grzadzinski, R., Carr, T., Colombi, C., McGuire, K., Dufek, S., Pickles, A., & Lord, C. (2016). Measuring changes in social communication behaviors: Preliminary development of the Brief Observation Of Social Communication Change (BOSCC). *Journal of Autism and Developmental Disorders, 46*(7), 2464–2479. https://doi.org/10.1007/s10803-016-2782-9

Grzadzinski, R., Janvier, D., & Kim, S. H. (2020). Recent developments in treatment outcome measures for young children with autism spectrum disorder (ASD). *Seminars in Pediatric Neurology, 34,* Article 100806. https://doi.org/10.1016/j.spen.2020.100806

Guttentag, S., Bishop, S., Doggett, R., Shalev, R., Kaplan, M., Dyson, M., Cohen, M., Lord, C., & Di Martino, A. (2022). The utility of parent-report screening tools in differentiating autism versus attention-deficit/hyperactivity disorder in school-age children. *Autism, 26*(2), 473–487. https://doi.org/10.1177/13623613211030071

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*(3), 325–342. https://doi.org/10.1007/s10803-005-0072-z

Havdahl, K. A., Bishop, S. L., Surén, P., Øyen, A.-S., Lord, C., Pickles, A., von Tetzchner, S., Schjølberg, S., Gunnes, N., Hornig, M., Lipkin, W. I., Susser, E., Bresnahan, M., Magnus, P., Stenberg, N., Reichborn-Kjennerud, T., & Stoltenberg, C. (2017). The influence of parental concern on the utility of autism diagnostic instruments. *Autism Research, 10*(10), 1672–1686. https://doi.org/10.1002/aur.1817

Hill, K. G., Woodward, D., Woelfel, T., Hawkins, J. D., & Green, S. (2016). Planning for long-term follow-up: Strategies learned from longitudinal studies. *Prevention Science, 17*(7), 806–818. https://doi.org/10.1007/s11121-015-0610-7

Hua, J., Li, Y., Ye, K., Ma, Y., Lin, S., Gu, G., & Du, W. (2019). The reliability and validity of Bayley-III cognitive scale in China's male and female children. *Early Human Development, 129,* 71–78. https://doi.org/10.1016/j.earlhumdev.2019.01.017

Jang, J., White, S. P., Esler, A. N., Kim, S. H., Klaiman, C., Megerian, J. T., Morse, A., Nadler, C., & Kanne, S. M. (2022). Diagnostic evaluations of autism spectrum disorder during the COVID-19 pandemic. *Journal of Autism and Developmental Disorders, 52*(2), 962–973. https://doi.org/10.1007/s10803-021-04960-7

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika, 39*(1), 31–36. https://doi.org/10.1007/BF02291575

Kanne, S. M., Mazurek, M. O., Sikora, D., Bellando, J., Branum-Martin, L., Handen, B., Katz, T., Freedman, B., Powell, M. P., & Warren, Z. (2014). The Autism Impact Measure (AIM): Initial development of a new tool for treatment outcome measurement. *Journal of Autism and Developmental Disorders, 44*(1), 168–179. https://doi.org/10.1007/s10803-013-1862-3

Kassambara, A. (2023). *rstatix: Pipe-friendly framework for basic statistical tests* (Version 0.7.2) [Computer software]. https://cran.r-project.org/web/packages/rstatix/index.html

Kat, S., Xu, L., Guo, Y., Ma, J., Ma, Z., Tang, X., Yang, Y., Wang, H., Li, X., & Liu, J. (2020). Reliability and validity of the simplified Chinese version of the aberrant behavior

checklist in chinese autism population. *Frontiers in Psychiatry, 11,* Article 545445. https://doi.org/10.3389/fpsyt.2020.545445

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2019). *lmerTest: Tests in linear mixed effects models* (Version 3.1-0) [Computer software]. https://CRAN.R-project.org/package=lmerTest

Lau, J. C. Y., Patel, S., Kang, X., Nayar, K., Martin, G. E., Choy, J., Wong, P. C. M., & Losh, M. (2022). Cross-linguistic patterns of speech prosodic differences in autism: A machine learning study. *PLOS ONE, 17*(6), Article e0269637. https://doi.org/10.1371/journal.pone.0269637

Law, T., Wei, Y., Wong, K. H. Y., Hu, Y., Tong, M. C. F., & Lee, K. Y. S. (2023). Feasibility and self-perceived effectiveness of an online training program on dysphagia in residential aged care homes in mainland China. *International Journal of Language & Communication Disorders, 58*(6), 1903–1911. https://doi.org/10.1111/1460-6984.12878

Lee, C. M., Altschuler, M. R., Esler, A. N., Burrows, C. A., & Hudock, R. L. (2023). Why are only some children with autism spectrum disorder misclassified by the Social Communication Questionnaire? An empirical investigation of individual differences in sensitivity and specificity in a clinic-referred sample. *Journal of Neurodevelopmental Disorders, 15*(1), Article 28. https://doi.org/10.1186/s11689-023-09497-7

Leng, L. L., Zhu, Y. W., & Zhou, L. G. (2023). Explaining differences in autism detection timing: Age of diagnosis and associated individual and socio-familial factors in Chinese children. *Autism, 28*(4), 896–907. https://doi.org/10.1177/13623613231187184

Lin, C.-S., Chang, S.-H., Cheng, S.-F., Chao, P.-C., & Chiu, C.-H. (2015). The preliminary analysis of the reliability and validity of the Chinese edition of the CSBS DP. *Research in Developmental Disabilities, 38,* 309–318. https://doi.org/10.1016/j.ridd.2014.12.023

Liu, S.-Y., Wang, X., Chen, Q., Chen, J.-J., Jin, C.-K., Zhan, X.-L., Guo, C.-H., Li, X.-H., Lin, L.-Z., & Jing, J. (2022). The validity and reliability of the simplified Chinese version of the Social Communication Questionnaire. *Autism Research, 15*(9), 1732–1741. https://doi.org/10.1002/aur.2726

Lord, C., Rutter, M., DiLavore, P., Risi, S., Gotham, K., & Bishop, S. (2012). *Autism Diagnostic Observation Schedule–Second Edition (ADOS-2)*. Western Psychological Corporation.

Maskey, R., Fei, J., & Nguyen, H.-O. (2018). Use of exploratory factor analysis in maritime research. *The Asian Journal of Shipping and Logistics, 34*(2), 91–111. https://doi.org/10.1016/j.ajsl.2018.06.006

McGillion, M., Pine, J. M., Herbert, J. S., & Matthews, D. (2017). A randomised controlled trial to test the effect of promoting caregiver contingent talk on language development in infants from diverse socioeconomic status backgrounds. *Journal of Child Psychology and Psychiatry, 58*(10), 1122–1131. https://doi.org/10.1111/jcpp.12725

Mullen, E. M. (1995). *Mullen Scales of Early Learning*. AGS.

Mundy, P., Block, J., Delgado, C., Pomares, Y., Van Hecke, A. V., & Parlade, M. V. (2007). Individual differences and the development of joint attention in infancy. *Child Development, 78*(3), 938–954. https://doi.org/10.1111/j.1467-8624.2007.01042.x

Naveed, S., Waqas, A., Amray, A. N., Memon, R. I., Javed, N., Tahir, M. A., Ghozy, S., Jahan, N., Khan, A. S., & Rahman, A. (2019). Implementation and effectiveness of non-specialist mediated interventions for children with autism spectrum disorder: A systematic review and meta-analysis. *PLOS ONE, 14*(11), Article e0224362. https://doi.org/10.1371/journal.pone.0224362

Nayar, K., Kang, X., Winston, M., Wong, P., & Losh, M. (2023). A cross-cultural study of visual attention in autism spectrum disorder. *Child Neuropsychology, 29*(3), 413–444. https://doi.org/10.1080/09297049.2022.2094904

Nayar, K., Kang, X., Xing, J., Gordon, P. C., Wong, P. C. M., & Losh, M. (2021). A cross-cultural study showing deficits in gaze-language coordination during rapid automatized naming among individuals with ASD. *Scientific Reports, 11*(1), Article 13401. https://doi.org/10.1038/s41598-021-91911-y

Oakland, T., & Harrison, P. L. (2011). *Adaptive Behavior Assessment System-II: Clinical use and interpretation*. Academic Press.

Pang, Y., Lee, C. M., Wright, M., Shen, J., Shen, B., & Bo, J. (2018). Challenges of case identification and diagnosis of autism spectrum disorders in China: A critical review of procedures, assessment, and diagnostic criteria. *Research in Autism Spectrum Disorders, 53,* 53–66. https://doi.org/10.1016/j.rasd.2018.06.003

Pepper, J., Weitzman, E., & Centre, H. (2004). *It takes two to talk: A practical guide for parents of children with language delays*. The Hanen Centre.

R Core Team. (2024). *R: A language and environment for statistical computing* (Version R Version 4.4.1) [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Reynolds, C. R. (2010). Behavior Assessment System for Children. In *The Corsini encyclopedia of psychology*. Wiley. https://doi.org/10.1002/9780470479216.corpsy0114

Rogers, S. J., Estes, A., Lord, C., Vismara, L., Winter, J., Fitzpatrick, A., Guo, M., & Dawson, G. (2012). Effects of a brief Early Start Denver model (ESDM)–based parent intervention on toddlers at risk for autism spectrum disorders: A randomized controlled trial. *Journal of the American Academy of Child & Adolescent Psychiatry, 51*(10), 1052–1065. https://doi.org/10.1016/j.jaac.2012.08.003

Rutter, M., Le Couteur, A., & Lord, C. (2003). *Autism Diagnostic Interview–Revised*. Western Psychological Services.

Sabri, S. (2013). Item analysis of student comprehensive test for research in teaching beginner string ensemble using model based teaching among music students in public universities. *International Journal of Education and Research, 1*(12).

Sparrow, S. S., Cicchetti, D. V., & Balla, D. (2005). *Vineland Adaptive Behavior Scales–Second Edition (Vineland-II)*. Pearson. https://doi.org/10.1037/t15164-000

Stahmer, A. C., Dababnah, S., & Rieth, S. R. (2019). Considerations in implementing evidence-based early autism spectrum disorder interventions in community settings. *Pediatric Medicine, 2,* Article 18. https://doi.org/10.21037/pm.2019.05.01

Tan, T. X., Yi, Z., Kamphaus, R. W., Wang, Y., Li, Z., & Cheng, K. (2020). Testing the reliability and construct validity of a chinese translation of BASC-3-SRP for 12- to 18-year-old youth. *Journal of Psychoeducational Assessment, 38*(5), 599–612. https://doi.org/10.1177/0734282919874257

Wetherby, A. M., & Prizant, B. M. (2002). *Communication and Symbolic Behavior Scales Developmental Profile–First Edition*. Brookes.

William, R. (2024). *psych: Procedures for psychological, psychometric, and personality research* (Version 2.4.6) [Computer software]. Northwestern University. https://CRAN.R-project.org/package=psych

Wu, J., & Pan, J. (2019). The scarcity of child psychiatrists in China. *The Lancet Psychiatry, 6*(4), 286–287. https://doi.org/10.1016/S2215-0366(19)30099-9

Wu, Z., Mak, M. C. K., Hu, B. Y., He, J., & Fan, X. (2019). A validation of the social skills domain of the Social Skills Improvement System–Rating Scales with Chinese preschoolers. *Psychology in the Schools, 56*(1), 126–147. https://doi.org/10.1002/pits.22193

Xinhua, C. (2023). *China's internet users total 1.079 bln with vibrant digital infrastructure, internet app growth*. The State Council, The People's Republic of China. https://english.www.gov.cn/archive/statistics/202308/28/content_WS64ec5126c6d0868f4e8dee21.html

Zhang, Z., Chen, X., & Qiu, C. (2020). *Vineland Adaptive Behavior Scales, Third Edition (Vineland-3) (Chinese version)*. Chinese Behavioral Science Corporation. https://www.mytest.com.tw/Infant_page.aspx?title=LF_Vineland3

Zhou, B., Zhou, H., Wu, L., Zou, X., Luo, X., Fombonne, E., Wang, Y., Yan, W., & Xu, X. (2017). Assessing the accuracy of the modified Chinese Autism Spectrum Rating Scale and Social Responsiveness Scale for screening autism spectrum disorder in Chinese children. *Neuroscience Bulletin, 33*(2), 168–174. https://doi.org/10.1007/s12264-017-0114-5

Zhou, R., Ma, B., & Wang, J. (2022). Difficulties in the diagnosis and treatment of children with autism spectrum disorder in China. *Journal of Autism and Developmental Disorders, 52*(2), 959–961. https://doi.org/10.1007/s10803-021-04997-8