



# Controlled Trial of a Short-term Intensive Parent Training Program within the Context of Routine Services for Autistic Children in China

Zuyi Fang<sup>a</sup>, Jamie M. Lachman<sup>b</sup>, Dongping Qiao<sup>a</sup>, and Jane Barlow<sup>b</sup>

<sup>a</sup>School of Social Development and Public Policy, Beijing Normal University, People's Republic of China;

<sup>b</sup>Department of Social Policy and Intervention, University of Oxford, United Kingdom

## ARTICLE INFO

Received 21 December 2021  
Accepted 18 April 2022

### Keywords:

Autism  
Parent training  
Real-world evaluation  
Propensity score  
Difference-in-differences

## ABSTRACT

Recent systematic reviews found limited rigorous research conducted to date of the effectiveness of parent training programs in reducing behavioral problems for autistic children in low- and middle-income countries. This study is aimed at evaluating the effectiveness of a short-term intensive parent training program for autistic children aged three to six in the context of routine service provision in China. A quasi-experiment was conducted involving the local implementing organization and using a waitlist control. Data were collected at baseline and immediate post-intervention. The primary outcome was child behavioral problems measured using the Child Behavior Checklist Externalizing scale. Between-group comparisons used a difference-in-differences design with propensity score weighting to reduce sources of bias. A process evaluation was undertaken in parallel to assess participant involvement, program acceptability, and delivery. The protocol was prospectively registered with ClinicalTrials.gov (NCT04257331). The final sample size was 111 (treatment: 63; comparison: 48). Results suggest that the program was associated with improvements in child externalizing behaviors ( $b = -2.71$ , 95% CI [-5.23, -0.18]), parental mental health symptoms ( $b = -5.96$ , 95% CI [-11.74, -0.17]), over-reactive parenting ( $b = -0.63$ , 95% CI [-0.98, -0.27]), and parental knowledge ( $b = 2.08$ , 95% CI [2.07, 2.17]). Exploratory analysis of factors related to implementation indicated that baseline parental mental health was related to participant engagement, and that satisfaction and engagement levels were potentially linked to positive treatment effects. Findings suggest that short-term intensive parent training programs that are provided by trained non-specialists, could potentially be used as an alternative to traditional prohibitively costly services that are delivered intensively for consecutive years in low-resource contexts. Follow-ups are needed to investigate its long-term benefits.

## Ensayo controlado de un programa de formación parental intensivo a corto plazo en el contexto de los servicios rutinarios para niños autistas en China

## RESUMEN

En revisiones sistemáticas recientes apenas se han encontrado investigaciones rigurosas que se hayan llevado a cabo hasta la fecha sobre la eficacia de los programas de formación parental para disminuir los problemas comportamentales de niños autistas en países de renta baja y media. En este estudio se pretende analizar la eficacia de un programa de formación parental intensivo a corto plazo para niños autistas de entre tres y seis años en el contexto de la prestación rutinaria de servicios en China. Se llevó a cabo un cuasiexperimento con la organización local a cargo de la aplicación y como control una lista de espera. Se recogieron datos de línea base e inmediatamente posteriores a la intervención. El resultado primario fueron los problemas comportamentales del niño medidos con la "Child Behavior Checklist Externalizing scale". Para las comparaciones entre grupos se utilizó un diseño de "diferencia en diferencias" con ponderación de la puntuación de propensión para disminuir las fuentes de error. En paralelo se llevó a cabo una evaluación de procesos para medir la implicación de los participantes, la aceptación del programa y su aplicación. El protocolo se registró prospectivamente en ClinicalTrials.gov (NCT04257331). La muestra final quedó constituida por 111 sujetos (de los cuales 63 constituían el grupo de tratamiento y 48 el de comparación). Los resultados indican que el programa se asociaba con la mejora de los comportamientos externalizadores infantiles ( $b = -2.71$ , 95% CI [-5.23, -0.18]), la salud mental parental ( $b = -5.96$ , 95% CI [-11.74, -0.17]), parentalidad sobrereactiva ( $b = -2.71$ , 95% CI [-5.23,

### Palabras clave:

Autismo  
Formación parental  
Evaluación del mundo real  
Puntuación de propensión  
Diferencia en diferencias

Cite this article as: Fang, Z., Lachman, J. M., Qiao, D., & Barlow, J. (2022). Controlled trial of a short-term intensive parent training program within the context of routine services for autistic children in China. *Psychosocial Intervention*, 31(2), 121-131. <https://doi.org/10.5093/pi2022a9>

Funding: ZF's involvement was partly funded by the Swire Charitable Trust. JML's involvement was supported by the University of Glasgow Social and Public Health Sciences Unit Complexity and Relationships in Health Improvement Programs of the Medical Research Council MRC UK and Chief Scientist Office (Grant: MC\_UU\_00022/1 and CSO SPHSU16, MC\_UU\_00022/3 and CSO SPHSU18). The funders played no role in the design, conduct, or interpretation of the analyses. Correspondence: [fangzuyi@bnu.edu.cn](mailto:fangzuyi@bnu.edu.cn) (Z. Fang).

ISSN:1132-0559/© 2022 Colegio Oficial de la Psicología de Madrid. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

-0.18]) y conocimientos parentales ( $b = 2.08$ , 95% CI [2.07, 2.17]). El análisis exploratorio de los factores relativos a la aplicación indicaba que la salud mental parental en la línea base se relacionaba con la implicación de los participantes y que el grado de satisfacción y de compromiso se relacionaba potencialmente con los efectos positivos del tratamiento. Los resultados indican que los programas de formación parental intensivos a corto plazo impartidos por personal no especialista entrenado podrían utilizarse potencialmente como alternativa a los servicios con un coste prohibitivo dispensados de modo intensivo en años consecutivos en contextos de recursos limitados. Se necesitan estudios de seguimiento para valorar sus ventajas a largo plazo.

Autism is characterized by impairments in reciprocal social-communicative interaction and restricted and repetitive behaviors or interests (American Psychiatric Association, 2013). Although the national prevalence of autism is yet to be estimated in China, a large regional study of children aged six to 12 years published in 2020 revealed a local rate of 0.70%, similar to the global estimate of 0.76% (Baxter et al., 2015; Zhou et al., 2020).

Autistic children experience more co-occurring behavioral problems (such as oppositional defiant behaviors and conduct problems) than their non-autistic cohorts (Simonoff et al., 2008; Totsika et al., 2011), which create further limitations in children's daily activities and impose more challenges to parenting (Hastings & Brown, 2002). Short-term parent training is widely used to support families of autistic children in low-resource settings. Due to the increasing number of diagnoses of autism, limited local service provision, and substantial economic cost of the autism-associated consequences at both family- and state-levels, such programs have been increasingly advocated globally as a cost-effective alternative to more conventional intensive interventions delivered for consecutive years (such as the Early Intensive Behavioral Intervention; Anan et al., 2008; Bearss et al., 2015; De Vries, 2016; Liao et al., 2022; Morris et al., 2011).

Short-term parent training programs may be a more scalable approach in Chinese settings. As with some other countries, access to autism treatments in China is prohibited by the widening service gap, absence of educational opportunities, high cost of private therapies, and insufficient financial support (Chang & Zaroff, 2017; Liao et al., 2020; Liao et al., 2022; P. Liu & Liu, 2018; McCabe, 2012). In a survey involving more than 3,800 Chinese parents of children with autism, over 80% of the respondents reported a lack of support and a strong need for parent training programs to improve parental efficacy and overcome a range of autism-induced challenges (Guo et al., 2014). Further, a qualitative study investigating the perspectives of caregivers and service providers with regard to a parent training program for autistic children in China, also emphasized the importance of delivering comprehensive program topics within a short period and providing substantial opportunities for practice and feedback, especially given the severe lack of parental knowledge about autism and a sense of urgency to intervene while their child was still young (Fang, Lachman, et al., 2022).

Five systematic reviews have been conducted to assess the evidence-base for parent training programs for autism (Fang, Barlow, et al., 2022; Nevill et al., 2018; Oono et al., 2013; Postorino et al., 2017; Strauss et al., 2013). The results showed that they can reduce child emotional and behavioral problems (which are not necessarily unique to autistic children) (Fang, Barlow, et al., 2022; Postorino et al., 2017), decrease autism symptoms (Nevill et al., 2018; Oono et al., 2013), promote child social communication and language development (Nevill et al., 2018; Strauss et al., 2013), and improve parent-child interaction (Fang, Barlow, et al., 2022; Oono et al., 2013). The review conducted by Fang, Barlow, et al. (2022) focused on programs delivered in China, by employing comprehensive searches in the English international and Chinese regional databases. Eleven short-term parent training programs for families of autistic children were identified, and it highlighted brevity as a common characteristic, despite the fact that families of autistic children in low-

resource settings often experience accumulated adversities and need substantial support, as well as an absence of rigorous evaluations of such programs being delivered in real-world settings (Fang, Barlow, et al., 2022).

This paper reports a controlled trial of a short-term intensive parent training program that has been supporting thousands of Chinese families of autistic children aged three to six years since the 1990s. The program is delivered by a long-established local NGO, known as Beijing Stars and Rain Education Institute for Autism (SREIA). Despite its long history and wide coverage, the program was only assessed using a single-group pre-post study (Xiong et al., 2010) and lacked evidence of effectiveness generated from a more robust experimental design. This report adheres to the Journal Article Reporting Standards (Appelbaum et al., 2018). Results of the evaluation might be used to inform service provision in other low- and middle-income countries (LMICs) where caregivers of autistic children face similar challenges.

## Method

### Study Design

We conducted a real-world quasi-experimental study to assess the effectiveness of the SREIA program in reducing child externalizing behaviors within the context of routine service provision. Participants were drawn from a permanent SREIA waiting list. Enrolment to the program was on a first-come, first-serve basis. Families attending between September 2020 and November 2020 comprised the treatment group. Those who signed up for the delivery between November 2020 and January 2021 comprised the waitlist control group. Due to COVID-19, three families on the waiting list decided not to attend the program and were thus also allocated into the control group. Outcomes were measured at baseline and immediate post-intervention. A concurrent process evaluation was also conducted to investigate program acceptability, participant involvement, and implementation.

### Participants

Recruitment of parent-child dyads was based on the standard criteria used by the SREIA program, including a) children between three to six years of age with an official diagnosis of autism issued by hospitals, and b) primary caregivers undertaking the responsibility for the daily childcare for at least five days a week. No restrictions are imposed on the severity of diagnosis or the level of child abilities, but as part of routine service, the program administers the Autism Treatment Evaluation Checklist (ATEC; Rimland & Edelson, 1999) at baseline and post-intervention time points to track child changes. The ATEC comprises four subscales, including Speech/Language/Communication, Sociability, Sensory/Cognitive Awareness, and Health/Physical/Behavior. It has a total score ranging from 0 to 179, with lower scores indicating fewer autism symptoms. However, it should be noted that the ATEC is mainly designed to capture treatment benefits and monitor child progress, rather than give a diagnosis or rate the severity. The decision not to administer other standardized measures in this research was based on concerns about participant

**Table 1.** Demographic characteristics

Demographic characteristics	Control ( <i>n</i> = 48)	Treatment ( <i>n</i> = 63)	Total ( <i>N</i> = 111)	<i>p</i> -value
Child gender, female	7 (14.58%)	11 (17.46%)	18 (16.22%)	.684
Child age, mean ( <i>SD</i> )	4.57 (1.21%)	4.80 (1.04)	4.70 (1.12)	.279
Ethnicity, minorities	3 (6.25%)	5 (7.94%)	8 (7.21%)	.734
Previous services, individual therapy	24 (50.00%)	40 (63.49%)	64 (57.66%)	.374
Medical history, Yes	20 (41.67%)	33 (52.38%)	53 (47.75%)	.229
Medication history, Yes	36 (75.00%)	45 (71.43%)	81 (72.97%)	.775
Comorbidity, Yes	29 (60.42%)	39 (61.90%)	68 (61.26%)	.790
Adult gender, female	42 (87.50%)	61 (96.83%)	103 (92.79%)	.060
Adult age, mean ( <i>SD</i> )	34.71 (5.18)	34.79 (6.28)	34.75 (5.80)	.942
Adult education, high school or less	26 (54.17%)	39 (61.90%)	65 (58.56%)	.511
Annual income, below the national average salary	11 (22.92%)	21 (33.33%)	32 (28.83%)	.230
Work status, full-time caregiver	27 (56.25%)	43 (68.25%)	70 (63.06%)	.194
Spouse work status, full-time caregiver	4 (8.33%)	6 (9.52%)	10 (9.01%)	.828
Household registration/Hukou, rural	22 (45.83%)	25 (39.68%)	47 (42.34%)	.631
Number of children, two or more	26 (54.17%)	30 (47.62%)	56 (50.45%)	.494

*p* < .05.

burden and overall data quality. Recruitment was conducted by the implementing organization inviting caregivers on the waiting list to the study. Each family had one parent-child dyad participating in the program. Financial reimbursements were provided to participants who completed data collection at both time points. Eight program practitioners were also recruited to the study.

Table 1 summarizes the demographic characteristics of the participants. More than 80% of the children involved in this program were male, in line with the fact that autism is about four times more common in boys than girls (Maenner et al., 2020). The mean age was 4.70 (*SD* = 1.12). The proportion of children from ethnic minority groups (7.21%, *n* = 8) was roughly in accordance with that reported in the Sixth National Population Census of China (8.49%) (National Bureau of Statistics of China, 2011). The ATEC total scores at baseline varied substantially, ranging from 16 to 118, with a mean score of 67.13 (*SD* = 21.29), speech/language/communication: range 0–26, *M* (*SD*) = 15.67 (6.49); sociability: 3–32, 15.99 (6.17); sensory/cognitive awareness: 5–31, 18.69 (5.79); health/physical/behavior: 2–41, 16.77 (8.15). Participating caregivers were predominantly female (92.79%, *n* = 103) aged 34.75 (*SD* = 5.8) on average. More than half were full-time caregivers (63.06%, *n* = 70). One tenth attended middle school or lower (17.12%, *n* = 19), 40% received high school-level education (41.44%, *n* = 46), and 40% attended college or more (41.44%, *n* = 46). About one third (28.83%, *n* = 32) had an annual income below the national average salary. Half of the families (50.45%, *n* = 56) had two or more children and slightly fewer have a rural household registration (42.34%, *n* = 47).

### Ethical Considerations and Preregistration

Ethical approval was granted by the University of Oxford (R67619/RE001) and Beijing Normal University (SSDPP-HSC2020001). Verbal informed consent was collected from all participants by trained research staff. The study was preregistered (NCT04257331).

### Intervention

The SREIA parent training program has a long history of providing services to families of autistic children in China. It was developed in the 1990s based on applied behavior analysis (ABA), which is grounded in operant conditioning (Cooper et al., 2007; Skinner, 1950), and has evolved to address issues identified during delivery of the program since then. It is aimed at reducing child behavioral

problems and promoting child development by increasing parental knowledge, improving parenting skills, promoting parental mental health, and creating social support. The program combines ABA-based protocols (such as discrete trial teaching) with social learning theory-based parenting techniques (such as instruction giving, ignore, and redirect) (Bandura, 1971) to improve child behavior and foster child skill development. Developmental perspectives are integrated by teaching parents to understand child developmental stages and domains and engage children in active parent-child communication and interaction, with the purpose of facilitating affective exchange and creating a foundation for learning other skills (Schreibman et al., 2015). Cognitive behavioral principles are also incorporated to increase parental knowledge and alter parental attitudes towards their child and themselves, as well as their situation and prospects, so that parental emotions and behaviors can be further improved. Additionally, a brief introduction to TEACCH [Treatment and Education of Autistic and related Communication-handicapped Children] strategies is provided to inform parents about the utilization of child visual strengths and the creation of structured work routines (Mesibov & Shea, 2010). Based on those underpinning theories, the program offers a wide range of techniques and is delivered five days a week for 11 weeks (Table 2).

Fourteen large didactic presentations on the core themes are offered across the 11 weeks and attended by all participating caregivers, with childcare provided by volunteers when there is no COVID-19 pandemic or restrictions on mass gatherings. Tailored training and practice opportunities are then arranged in smaller groups comprising ten parent-child dyads to improve the use of skills. The program is primarily delivered in-person, while the fourteen presentations were also streamed online to reduce mass gathering during COVID-19. Individual pre-consultations are offered with the aim of enabling program practitioners to better understand family strengths and concerns. An individualized treatment plan is created collaboratively and followed for each dyad, covering targeted child behaviors, caregiver areas for improvement, and other priorities identified by caregivers. Materials and handouts are provided. Practitioners are from various backgrounds (without healthcare-related education) but all trained and certified by the implementing organization. To promote participant engagement and learning, the program integrated a variety of delivery techniques and substantial opportunities for practice (Table 2).

The program has two delivery formats, which are implemented concurrently and share the same program content, with the purpose of exploring ways to reduce cost while improving outcomes. Both

**Table 2.** SREIA Program Topics and Delivery Techniques

Topics	Delivery Techniques
Knowledge of ASD and child development	Individual consultation
Child behavioural adjustment and behaviour management techniques based on ABA, including reinforcement (praise & rewards), prompts, fading, shaping, and chaining	Didactic presentation
TEACCH – “Structured TEACCHing”	Practitioner observation
Instruction giving	Facilitator modelling new techniques
Ignore	Cases and stories
Redirect	Practice in sessions and in-vivo feedback
Parent-child communication and interaction (including parent-child play)	Role play
Emotional labelling	Homework and feedback
Fine and gross motor skills	Videorecording and feedback
Communicate with family members	Group discussion on skills and experiences
Parental self-regulation	Group problem solving
Parental attitude and expectation	Physical activity
Skill maintenance and generalisation	Reinforcement and active listening by practitioners
Plan for the future	Online group communication
	Informal peer support group
	Printed materials
	Childcare

versions divide participants to several groups based on children's ATEC assessments. Version A has four groups of dyads who attend didactic sessions together in a larger group but practice sessions separately in smaller groups, whereas Version B has three groups of dyads and delivers all (didactic and practice) sessions separately for each group. Participants are offered the option of selecting either version when signing up for the waiting list.

### Power Calculation

As a real-world evaluation conducted in a routine service context, the sample size was determined by the service capacity of the implementing organization. The minimal detectable standardized mean difference from two-sample, two-tailed *t*-test was -0.57 on the primary outcome measure. This was based on a final sample size of 111 with a significance level of .05 and statistical power of .80.

### Outcomes and Measures

The outcomes of interest were identified by the program practitioners and developers in discussion with the research team. Child behavioral problems was considered to be the primary outcome reflecting the wider range of behavioral challenges experienced by parents on a daily basis and that the program aimed to address. It was assessed using the Externalizing scale of the Child Behavior Checklist (CBCL) for ages 1.5-5 (25 items) (Achenbach & Rescorla, 2001), as the majority of children were under six and similar levels of internal consistency of CBCL 1.5-5 was shown in autistic children aged five and those aged six (Basten et al., 2014). The Externalizing scale contains two subscales: Attention Problems and Aggressive Behaviours. The Attention Problems subscale consists of five items (e.g., “can't concentrate”, “quickly shifts”, and “wanders away”), and the Aggressive Behaviours subscale has 19 items (e.g., “can't stand waiting” and “demands met”). The scale has been applied in children with a range of developmental disabilities, including autism, in high-income countries, such as the Incredible Years-Autism Spectrum and Language Delays Programme in the UK (McIntyre, 2008; Williams et al., 2020) and the Parents' Plus Programme for children with developmental disabilities in Ireland (Quinn et al., 2007). It has demonstrated good cross-cultural factorial validity in China and adequate internal consistency in autistic children (J. Liu et al., 2011; Pandolfi et al., 2009).

Impact on several secondary outcomes was examined. “Autism symptoms” was assessed using the Chinese version of the ATEC (Rimland & Edelson, 1999). “Parenting style” was measured using the Over-Reactivity subscale of the Arnold-O'Leary Parenting Scale (Arnold et al., 1993) and the Supporting Positive Behavior subscale of the Parenting Young Children Scale (McEachern et al., 2012). “Parental mental health” was assessed using the Chinese version of the Depression Anxiety Stress Scale (Lovibond & Lovibond, 1995). “Family function” was assessed using the Chinese version of the Family APGAR scale (Smilkstein et al., 1982). All scales have been validated in China, except the Over-Reactivity subscale, which has been validated in families with high risks for child behavioral problems in the U.S. “Parental knowledge” was measured using a questionnaire designed for the SREIA evaluation to assess parental understanding of autism and child behaviors. For questionnaires that do not have a Chinese version, the translation followed the WHO guideline of the Process of Translation and Adaptation of Instruments (World Health Organization, 2020). All outcome measures were reported by participating parents.

“Program acceptability” was assessed only at post-intervention using a satisfaction questionnaire tailored to the SREIA program and the Chinese context (Incredible Years, 2020). The level of “participant involvement” was captured by calculating rates of attendance, engagement, and program dropout. The attendance and engagement rates were recorded using an attendance and engagement registry, reported by practitioners each day. The level of caregiver engagement was judged based on four criteria: responsiveness to practitioners, interactions with other caregivers, participation in group discussion, and completion of homework activities. Program dropout rate was calculated based on families who missed at least ten days. “Implementation” was measured through fidelity and quality of delivery. Fidelity was assessed using practitioner checklists. Practitioners reported whether they delivered the core components prescribed in the manual each week. The fidelity score comprised the overall ratio of delivered components to prescribed components. Quality of delivery was assessed by research staff observing sessions and filling out a quality of delivery survey. The survey contains 24 items, assessing the general skills of facilitating parent groups (such as knowledge of content, teaching skills, sensitivity to participants' feelings and experiences, ability to handle tension, etc.) and skills of modelling parenting techniques and facilitating practice sessions (such as appropriately setting up the space and providing adequate



support during participant practice). A total score is summed. Two sessions for each practitioner were randomly selected for research staff to rate the quality of delivery. The program process was also investigated using qualitative in-depth interviews and focus group discussions, the results of which are reported separately (Fang, Lachman, et al., 2022).

### Data Collection and Analysis

Data collection took place between August to September 2020 for the baseline assessment, and November to December 2020 for the immediate post-test. Trained researchers administered the parent questionnaires to the treatment and control groups, by reading out the questions to caregivers. All data collectors had a master's degree in social work and were trained in a range of relevant topics. Caregivers were requested to keep their allocation condition private from outcome data collectors and the data collectors were asked whether they knew to which group the participant belonged after each interview. The process evaluation data, including satisfaction survey, were collected by a separate group of research staff. No violation of blinding was detected.

Data analyses were conducted in R v4.0.3 and followed the principles of intent-to-treat by including all participant recruited at baseline. Baseline group differences were explored using two independent samples *t*-tests and chi-square tests. To reduce selection bias and improve baseline balance for the between-group comparison, the treatment effects were estimated using a difference-in-differences (DiD) design with propensity score weighting (PSW) (Stuart et al., 2014). Propensity scores were calculated using logistic regression. The key covariates were selected based on the relevant literature and included the baseline outcome measures and child gender. The treatment and control groups were weighted using the overlap weighting scheme to avoid extreme weights and ad hoc trimming (Li & Thomas, 2019). Group balance was checked using the absolute standardized mean difference, with a rule of thumb that the absolute standardized mean difference should be controlled under 0.1 (Rubin, 2001).

Weighted DiD models were then fitted, with accounting for the key covariates, to further adjust for unobserved variables and secular time trends (Wing et al., 2018). Dependency of outcomes for participants within each version and each caregiver group were checked using intra-cluster correlation coefficients and, when necessary, accounted for in linear mixed models with caregiver group or delivery version being treated as the random effect. Exploratory analyses were also conducted by replacing the binary treatment variable with a three-level variable in the regression models to assess the differences in treatment effects between the two versions of delivery. Robustness checks were conducted using DiD with inverse probability weighting and kernel matching (Austin & Stuart, 2015; Caliendo & Kopeinig, 2008). The ignorability assumption of propensity score weighting and parallel trend assumption of DiD were checked by examining the effect of group assignment on the baseline primary outcome measure and treating financial health (a placebo outcome associated with child externalizing behaviors (Slopen et al., 2010) but not affected by group assignment) as the outcome variable in the weighted DiD (Imbens, 2004; Wing et al., 2018). Beta coefficients were converted to Cohen's *d* for reporting using the method suggested by Brysbaert and Stevens (2018).

Multivariate models were created to examine potential factors associated with participant involvement, with accounting for confounders drawn from existing evidence, including quality of delivery, baseline child externalizing problems, parental mental health, autism symptoms, parental education, adult age, adult gender, and family income. The possible associations between implementation factors and the treatment effect were investigated

by first using univariate models and then adjusting for the key covariates. The missing data comprised less than 5% of the entire dataset and were compensated using REML (Hartung et al., 2008). No data were missing in variables that were included in the DiD models. Residuals in all regression models were checked for normality and outliers.

**Transparency and openness.** We reported how we determined our sample size, all manipulations, and all measures in the study, and we followed the JARS guidelines (Appelbaum et al., 2018). We excluded data from three participants, due to misdiagnosis and data withdrawal. The anonymized individual participant data, analytic code, unpublished full study protocol, and other research materials will be available upon request from researchers who provide a methodologically sound proposal. To gain access, requestors will need to sign a data sharing agreement.

**Community involvement.** The implementation organization took part in the planning of the overall research project, by contributing to the development of the research design, research questions, and evaluation tools. The recruitment of participants and implementation of the program were undertaken by the organization using its existing resources. The organization did not participate in data collection or analysis but has been actively involved in the interpretation of results and taking the lead in conveying the research findings to a wider autism community in China.

#### Recruitment and Baseline Assessment

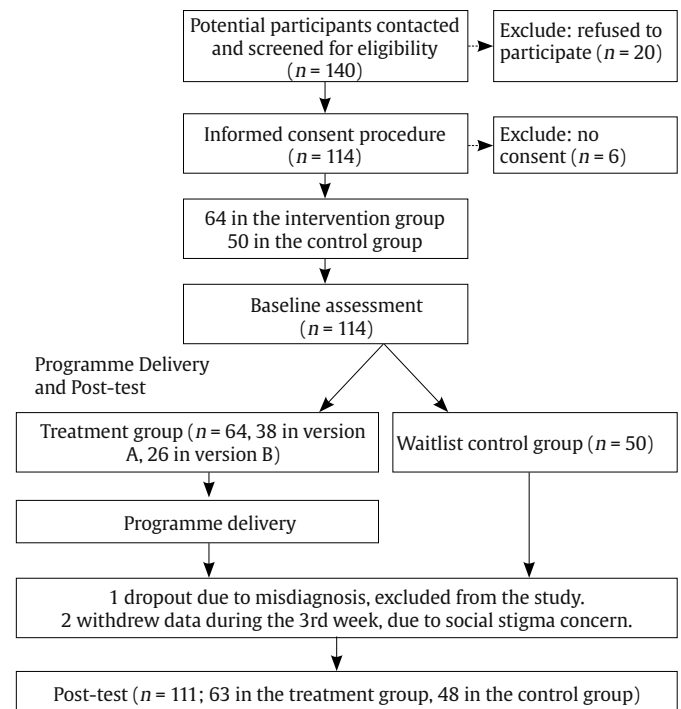


Figure 1. Participant Flow Chart.

### Results

Figure 1 shows the flow of participants through the study. The final sample involved 63 caregiver-child dyads in the treatment group—37 in Version A and 26 in Version B—and 48 in the control group. Table 3 presents the absolute standardized mean differences for each baseline measure. Table 4 presents the baseline and post-intervention scores for all outcome measures. While the *t*-tests

and chi-square tests detected no significant group difference in demographic characteristics (Table 1) and outcome measure at baseline, several key covariates demonstrated substantial absolute standardized mean differences. The overlap weighting successfully reduced the absolute standardized mean differences across all covariates and achieved excellent group balance.

**Table 3.** Absolute Standardized Mean Differences (ASMD)

Key Covariates	Pre-Weighting ASMD	Post-Weighting ASMD
Baseline child externalizing behaviors	-0.21291	-2.19E-13
Baseline caregiver mental health	-0.03578	-2.43E-14
Baseline parental over-reactivity	-0.15566	-1.03E-13
Baseline positive parenting	0.331887	2.76E-13
Baseline autism symptoms	-0.03055	-4.95E-14
Baseline parental knowledge	-0.04685	-5.89E-14
Baseline family function	0.142515	2.28E-13
Child gender	0.077777	7.38E-14

### Estimation of Treatment Effects

The estimations of DiD are summarized in Table 5. Analyses found that the program was linked to reduced child externalizing problems ( $b = -2.71$ , 95% CI [-5.23, -0.18], Cohen's  $d = -2.80$ ), compared to the comparison group. The results of robustness checks (inverse probability weighting:  $b = -2.76$ , 95% CI [-5.08, -0.44]; kernel matching:  $b = -3.28$ , 95% CI [-4.59, -1.97]) were in line with the estimation and did not detect violation of underlying assumptions.

In terms of secondary outcomes, analyses found that the program was associated with increased parental knowledge of autism and child development ( $b = 2.08$ , 95% CI [2.07, 2.17], Cohen's  $d = 2.91$ ), improved parental mental health ( $b = -5.96$ , 95% CI [-11.74, -0.17],

Cohen's  $d = 2.89$ ), and reduced over-reactive parenting ( $b = -0.63$ , 95% CI [-0.98, -0.27], Cohen's  $d = -5.02$ ), in comparison to control condition. No significant associations were detected for autism symptoms ( $b = -1.53$ , 95% CI [-6.44, 3.38], Cohen's  $d = -0.86$ ), family function ( $b = -0.08$ , 95% CI [-0.88, 0.71], Cohen's  $d = -0.28$ ), or positive parenting practices ( $b = 0.16$ , 95% CI [-1.89, 2.21], Cohen's  $d = 0.22$ ).

The exploratory analysis on version difference indicated that, compared to the comparison group, Version B was linked to a reduction in externalizing problems ( $b = -3.77$ , 95% CI [-6.26, -1.28], Cohen's  $d = -4.26$ ) and an improvement in parental mental health ( $b = -12.52$ , 95% CI [-19.87, -5.18], Cohen's  $d = -4.79$ ), whereas Version A did not differ from the control group (externalizing problems:  $b = -2.08$ , 95% CI [-4.56, 0.41], Cohen's  $d = -2.35$ ; mental health:  $b = -1.02$ , 95% CI [-8.35, 6.31], Cohen's  $d = -0.39$ ). Both versions were associated with an increase in parental knowledge (VA:  $b = 1.49$ , 95% CI [0.28, 2.69], Cohen's  $d = 3.47$ ; VB:  $b = 3.03$ , 95% CI [1.82, 4.24], Cohen's  $d = 7.05$ ) and a reduction in parental over-reactivity (VA:  $b = -0.72$ , 95% CI [-1.15, -0.28], Cohen's  $d = -4.68$ ; VB:  $b = -0.44$ , 95% CI [-0.88, -0.01], Cohen's  $d = -2.86$ ). However, due to the small subgroup sample sizes, the analyses were underpowered.

### Implementation Evaluation

Table 6 shows that the level of participant involvement was high, with most caregivers attending 98% of the time and the average engagement score being 5.58 ( $SD = 0.32$ ). The main reasons for absence were child illness and family issues. On average, 95% of pre-specified topics and activities were implemented across the seven caregiver groups based on the fidelity checklists, with a mean quality of delivery score of 51.86 ( $SD = 1.13$ ). The overall satisfaction achieved 78.19 ( $SD = 6.94$ ).

**Table 4.** Baseline and Post-intervention Outcome Measures

Outcome Measures	Control ( <i>n</i> = 48)		Treatment ( <i>n</i> = 63)						<i>p</i> -value <sup>1</sup>
	Baseline <i>M</i> ( <i>SD</i> )	Post-test <i>M</i> ( <i>SD</i> )	Treatment ( <i>n</i> = 63)		Version A ( <i>n</i> = 37)		Version B ( <i>n</i> = 26)		
			Baseline <i>M</i> ( <i>SD</i> )	Post-test <i>M</i> ( <i>SD</i> )	Baseline <i>M</i> ( <i>SD</i> )	Post-test <i>M</i> ( <i>SD</i> )	Baseline <i>M</i> ( <i>SD</i> )	Post-test <i>M</i> ( <i>SD</i> )	
Primary Outcome									
Child Externalizing Behavior (CBCL)	19.79 (6.37)	19.75 (6.02)	21.30 (7.74)	17.87 (6.51)	20.54 (8.21)	18.08 (7.07)	22.38 (7.03)	17.58 (5.74)	.275
Secondary outcomes									
Autism Symptoms (ATEC)	66.75 (23.81)	62.21 (26.39)	67.41 (19.35)	60.92 (17.82)	67.24 (19.93)	61.73 (19.06)	67.65 (18.87)	59.77 (16.19)	.872
Parental Knowledge	26.96 (3.93)	27.71 (3.85)	27.13 (3.24)	29.71 (2.78)	27.54 (2.92)	29.16 (2.64)	26.54 (3.61)	30.50 (2.83)	.805
Parental Mental Health (DASS)	34.83 (24.16)	31.00 (24.81)	35.65 (21.45)	25.65 (18.63)	35.35 (24.41)	29.68 (22.26)	36.08 (16.80)	19.92 (9.46)	.851
Family Function (APGAR)	6.27 (2.52)	6.81 (2.52)	5.89 (2.83)	6.68 (2.48)	6.35 (2.85)	6.68 (2.51)	5.23 (2.73)	6.69 (2.49)	.462
Positive Parenting (PARYC)	35.29 (5.66)	35.50 (6.88)	33.37 (5.95)	34.90 (4.97)	33.30 (5.66)	35.59 (4.56)	33.46 (6.46)	33.92 (5.43)	.087
Parental Over-Reactivity (PS)	3.43 (1.16)	3.45 (1.29)	3.61 (1.20)	2.98 (0.97)	3.72 (1.27)	2.83 (0.98)	3.46 (1.09)	3.19 (0.94)	.419

Note. <sup>1</sup>Baseline differences between the treatment and control groups.

**Table 5.** Difference-in-differences Model Estimates

Outcome Measures	Difference-in-differences (treatment vs. control) $b$ [95% CI]	Cohen's $d$	Difference-in-differences (version A vs. control) $b$ [95% CI]	Cohen's $d$	Difference-in-differences (version B vs. control) $b$ [95% CI]	Cohen's $d$
<b>Primary outcome</b>						
Child Externalising Behaviour (CBCL)	-2.71* [-5.23, -0.18]	-2.80	-2.08 [-4.56, 0.41]	-2.35	-3.77** [-6.26, -1.28]	-4.26
<b>Secondary outcomes</b>						
ASD Symptoms (ATEC)	-1.53 [-6.44, 3.38]	-0.86	-0.82 [-7.13, 5.49]	-0.36	-3.40 [-9.72, 2.93]	-1.51
Parental Knowledge	2.08* [2.07, 2.17]	2.91	1.49* [0.28, 2.69]	3.47	3.03*** [1.82, 4.24]	7.05
Parental Mental Health (DASS)	-5.96* [-11.74, -0.17]	-2.89	-1.02 [-8.35, 6.31]	-0.39	-12.52** [-19.87, -5.18]	-4.79
Family Function (APGAR)	-0.08 [-0.88, 0.71]	-0.28	-0.34 [-1.35, 0.66]	-0.95	0.24 [-0.77, 1.24]	0.67
Positive Parenting (PARYC)	0.16 [-1.89, 2.21]	0.22	1.16 [-1.30, 3.62]	1.32	-0.61 [-3.08, 1.85]	-0.70
Parental Over-Reactivity (PS)	-0.63*** [-0.98, -0.27]	-5.02	-0.72** [-1.15, -0.28]	-4.68	-0.44* [-0.88, -0.01]	-2.86

\* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$ .

**Table 6.** Process Measures, Mean (SD)

	Total (N = 63)	Version A (n = 37)	Version B (n = 26)
Engagement score (maximum 6)	5.58 (0.32)	5.65 (0.32)	5.48 (0.31)
Attendance rate (maximum 1)	0.98 (0.03)	0.99 (0.03)	0.98 (0.03)
Satisfaction score (maximum 84)	78.19 (6.94)	77.49 (7.65)	79.19 (5.80)
Quality of delivery (maximum 72)	51.86 (1.13)	51.76 (0.25)	52.01 (1.74)
Fidelity rate (maximum 1)	0.95 (0.04)	0.98 (0.01)	0.91 (0.04)

The multivariate models showed that more parental mental health problems at baseline were related to lower engagement levels ( $b = -0.003$ , 95% CI [-0.006, -0.001]) but not attendance rate. No other factors examined were significantly associated with participant involvement (Table 7).

**Table 7.** Multivariate Models: Factors Associated with Participant Engagement and Attendance

	Model 1: Engagement	Model 2: Attendance
Baseline child behavior	0.002 [-0.006, 0.009]	0.001 [-0.000, 0.002]
Baseline parental mental health	-0.003* [-0.006, -0.001]	0.000 [-0.001, 0.000]
Baseline autism symptoms	-0.003 [-0.006, 0.000]	0.000 [-0.001, 0.000]
Quality of delivery	-0.107 [-0.246, 0.032]	-0.002 [-0.009, 0.004]
Adult education	-0.029 [-0.142, 0.085]	-0.003 [-0.021, 0.015]
Annual income	0.044 [-0.080, 0.165]	0.008 [-0.011, 0.027]
Adult age	-0.012 [-0.024, 0.000]	-0.001 [-0.003, 0.001]
Adult gender	0.321 [-0.080, 0.712]	0.013 [-0.048, 0.074]

\* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$ .

Further analysis conducted in the treatment group found that higher attendance rates were linked to fewer externalizing problems ( $b = -66.60$ , 95% CI [-125.30, -7.80]). While this relationship was not present after accounting for key covariates, higher satisfaction and engagement levels were both associated with better

treatment effects in the treatment group (satisfaction:  $b = -0.24$ , 95% CI [-0.44, -0.03]; engagement:  $b = -6.81$ , 95% CI [-12.19, -1.44]) (Table 8).

## Discussion

To our knowledge, this study is the first to evaluate a short-term intensive parent training program with a long history of supporting families of autistic children in China using a controlled trial. The evaluation was conducted in a routine service context with substantial effort to reduce sources of bias using a difference-in-differences design with propensity score weighting. Findings of this study might inform service provision in other LMICs where there is also a severe lack of resources for families of autistic children.

## Program Effects

The results showed that the program was associated with a reduction in child externalizing behaviors, a change that equates to a Cohen's  $d$  of -2.80, which is generally consistent with the finding (Hedges'  $g = -1.47$ ) of a previous systematic review of such programs in China (Fang, Barlow, et al., 2022). Although the systematic review reported Hedges'  $g$  rather than Cohen's  $d$ , research indicates that they produce extremely similar estimates when the sample size is larger than 20 (Lin & Aloe, 2021). However, our effect size was larger than other existing studies synthesizing trials conducted primarily in high-income countries and yielding effect sizes ranging from SMD = -0.39 to -0.59 (Postorino et al., 2017; Skotarczak & Lee, 2015; Tellegen & Sanders, 2013). There may be multiple reasons for this. First, the program provided substantial opportunities for participants to observe practitioner modelling, practice in sessions, and receive in vivo feedback, which promoted caregiver acquisition of skills. In addition, the wider evidence shows that parent training programs typically have a greater impact for families living in low-resource settings and who have an urgent need to address child behavior problems (Gardner et al., 2015; Leijten et al., 2013). The Chinese context is characterized by a severe lack of autism-related resources, and caregivers were in desperate need of evidence-

**Table 8.** Relationships between Implementation Factors and Child Externalizing Behaviors Post-intervention

Variable	Univariate Models				
	Model 1	Model 2	Model 3	Model 4	Model 5
Quality of delivery	0.25 [-1.20, 1.70]				
Satisfaction		-0.20 [-0.44, 0.03]			
Engagement score			-1.79 [-6.90, 3.30]		
Attendance rate				-66.60* [-125.30, -7.80]	
Fidelity rate					22.3 [-18.00, 62.00]
Variable	Multivariate Models				
	Model 1	Model 2	Model 3	Model 4	Model 5
Quality of delivery	0.03 [-1.29, 1.36]				
Satisfaction		-0.24* [-0.44, -0.03]			
Engagement score			-6.81* [-12.19, -1.44]		
Attendance rate				-43.78 [-98.35, 10.79]	
Fidelity rate					15.88 [-18.43, 50.18]
Baseline child behavior	-0.51*** [-0.72, -0.30]	-0.55*** [-0.75, -0.34]	-0.50*** [-0.70, -0.30]	-0.47*** [-0.68, -0.27]	-0.50*** [-0.71, -0.29]
Baseline autism symptoms	0.01 [-0.08, 0.11]	0.02 [-0.07, 0.11]	-0.02 [-0.12, 0.07]	0.01 [-0.08, 0.10]	0.01 [-0.08, 0.10]
Baseline parental mental health	0.02 [-0.06, 0.10]	0.01 [-0.07, 0.09]	-0.02 [-0.10, 0.06]	0.00 [-0.07, 0.08]	0.02 [-0.06, 0.10]
Baseline parental overreactivity	0.38 [-0.88, 1.65]	0.46 [-0.75, 1.67]	1.02 [-0.28, 2.31]	0.54 [-0.71, 1.79]	0.36 [-0.90, 1.61]
Baseline positive parenting	0.11 [-0.18, 0.39]	0.06 [-0.21, 0.33]	0.15 [-0.12, 0.42]	0.11 [-0.16, 0.39]	0.11 [-0.16, 0.39]
Baseline parental knowledge	0.18 [-0.31, 0.66]	0.08 [-0.38, 0.55]	0.20 [-0.26, 0.65]	0.14 [-0.33, 0.61]	0.16 [-0.32, 0.64]
Baseline family function	-0.18 [-0.73, 0.37]	-0.12 [-0.64, 0.04]	-0.34 [-0.87, 0.19]	-0.24 [-0.77, 0.30]	-0.18 [-0.72, 0.36]
Child gender	0.73 [-3.23, 4.69]	1.05 [-2.60, 4.71]	-0.66 [-4.44, 3.12]	0.26 [-3.52, 4.05]	0.94 [-2.88, 4.75]

\* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$ .

based techniques with which to tackle some of the problems being experienced by their children, which could have increased the effect sizes. Autistic children demonstrate more behavior problems than children without autism or with other developmental disabilities (Totsika et al., 2011), and the wider research suggests that programs targeting families of children with existing behavior problems have stronger effects than universal or selective programs (Leijten et al., 2019). While the severity of the problems in children involved in this study may also be reflected by caregivers' determination to participate in the program amidst all the challenges caused by COVID19, this might also have introduced a potential selection bias, in terms of making the sample less representative.

Analyses of secondary outcomes indicated that the program was also linked to better parental mental health, and a reduction in over-reactive parenting. This result differs from that of a previous review that found no group difference in parental stress in a parent-mediated early intervention delivered to caregivers of autistic children aged 17-72 months (Oono et al., 2013). This discrepancy may be because caregivers in the current study experienced an inclusive environment for the first time, which empowered and guided them towards greater acceptance of their children and themselves. Also, the improved parental self-efficacy and their perceived progress in child development might have created a sense of hope and increased feelings of control.

While we found no evidence of benefit in terms of reducing autism symptoms, which again is not consistent with the findings of the earlier reviews (Nevill et al., 2018; Oono et al., 2013), this may reflect the challenges of introducing a large number of techniques within a short timeframe, the absence of a long-term follow-up period, and the fact that it takes time for the techniques to become established (van Aar et al., 2017). The lack of association with parenting practices could be attributed to the services and information previously obtained by the participants, which predominantly emphasized the use of positive reinforcements. In addition, in order to reduce participant burden and increase the overall quality of data, a short scale was used to assess positive parenting practice, leading to a trade-off with regard to scale sensitivity. The absence of an impact on family functioning may be due to the same reason. Further, consistent with earlier research, the data suggests that conflicts occurred when non-participating family members did not understand the diagnosis or new parenting techniques (Mockford & Barlow, 2004).

The exploration of version differences indicated that better practitioner-dyads ratio was related to fewer child externalizing behaviors and better parental mental health, possibly as a result of the more individualized feedback and reduced practitioner workload. However, the results should be interpreted with caution due to the small samples available for subgroup analysis.

## Implementation Evaluation

Although previous evidence shows that engaging parents can be a challenge for providers of parenting interventions (Axford et al., 2012), this program was delivered with a high attendance and engagement rate, in spite of its high intensity of five days a week over 11 consecutive weeks. The reasons for this may be three-fold. First, the substantial service gap and prohibitive cost of prolonged treatments might have led to a degree of urgency in terms of participant needs for parent training and social support. Second, the participants may have been highly motivated because the period from two to six years of age was considered by them to be the most 'critical' and the only chance for intervention, especially given the absence of local services for older autistic children (Chang & Zaroff, 2017). Third, a variety of delivery techniques were used to promote engagement, such as ongoing and flexible communication, reinforcement by practitioners, and peer support. In terms of quality of delivery, the program was

rated as medium, which might be due to the heavy workload and lack of organizational support for practitioners (Fang, Lachman, et al., 2022).

The exploratory analyses to examine factors affecting participant involvement showed that a lower engagement rate was related to more baseline mental health symptoms, which is consistent with findings from previous research (Carr et al., 2016). While the literature tends to suggest that families from disadvantaged backgrounds may have lower attendance and engagement rates (Chacko et al., 2016), the current study found no such relationship, possibly reflecting the dearth of services and the need for support across the different socio-economic groups in LIMC contexts.

The analyses showed that a higher level of program satisfaction and engagement appeared to be associated with improved child behavioral adjustment, whereas the quality of delivery, attendance rate, and fidelity score made little difference. However, the results should be interpreted with caution, as the estimation of dose effect was limited by the small number of parent groups and the similarly high attendance and fidelity rates across groups. Future research should explore further the relationships between dosage and outcomes using methods such as complier average causal effect.

## Strengths and Limitations

This study has several limitations. First, as a real-world evaluation of a program that is embedded within routine services, it was not possible to conduct randomization, which could have resulted in unidentified confounders. Also, some families withdrew due to COVID-19, potentially increasing the risk of bias. Second, it was not possible to conduct multiple assessments prior to the program to test the underlying parallel trends assumption of DiD due to COVID-19, again pointing to the risk of unobserved confounders. To mitigate the impact of bias, key covariates identified from the literature were included. The robustness checks also showed that unobserved factors were time-invariant and did not affect the estimation of outcomes (Wing et al., 2018). In addition, although this study involved more participants than other existing evaluations in China (Fang, Barlow, et al., 2022), the sample size was still small, limiting the power and precision to identify treatment effects. Further, the outcome measures relied on parent-report data, which may be subject to social desirability bias, and we were not able to triangulate the data using practitioner-report or observational measures, due to resource constraints. Efforts were, however, made to reduce this source of bias, by using measurement tools with adequate reliability and blinding data collectors to group allocation. The measurement of participant engagement relied on practitioner report. A consultation on criteria for the level of engagement was conducted with the practitioners to complement findings from the literature. However, the reliability of the results could have been further improved by rating the accuracy of parents' practice with children and using standardized questionnaires. The study had limited male caregivers, undermining the generalizability of the results. Last, the data collection time point was immediately post-intervention. Longer follow-up periods are necessary to examine what proportion of techniques become firmly established and the extent to which the effects are maintained over time.

Despite the limitations, the study has a number of strengths. It is one of the first real-world evaluations of a parent training program for autistic children conducted in a routine service context, thereby being more likely to provide valuable information about the range of participants in clinical practice. Local stakeholders being actively engaged in the research also ensured that the research was locally relevant and increased the possibility of evidence uptake by community advocates and policymakers. Moreover, while many autism treatments have low levels of diversity in terms of the



participants, being delivered either to families with high or low socio-economic status, this program was attended by participants from diverse backgrounds (e.g., 59% without a college degree and 42% with a rural household registration) and children with different levels of autism-related challenges, which implies that the program may be well applicable to a wide range of families. In addition, previous reviews of parent training for autism commonly highlight the inadequate reporting of their implementation and the potential impact of this. The concurrent process evaluation within this study provided important information with regard to program implementation while also allowing us to integrate the process data with outcome assessment to examine the associations between treatment effects and implementation variability (Oakley et al., 2006). Also, the DiD method was combined with propensity score weighting to compensate for the lack of randomization in the context of routine service and remove secular trend effects.

## Implication and Conclusion

The findings of this study provide tentative support for the use of short-term intensive and practice-based parent training programs for families of young autistic children in LMICs to promote child behavioral adjustment and reduce wider health disparities. Although treatment models for autism can entail a variety of techniques and formats, researchers commonly advocate for intensive interventions that begin at an early age (National Research Council, 2001). Early intensive behavioral interventions have been found to be effective in promoting child development (Reichow et al., 2018), but can often be inaccessible due to the high cost and requirement for qualified professionals. As such, short-term parent training and parent-directed treatments have been developed as an alternative for families of young autistic children. The high participant commitment and satisfaction in the current study possibly reflects the substantial need for such programs in areas where there is a severe lack of services. Program providers were from a range of disciplines and trained by the implementing agency, thereby demonstrating the possibility of having non-specialists deliver the programs following training. Furthermore, as caregivers often play a key role as the first advocate and support for autistic children, such parent training programs would not only increase access to services, but also enable caregivers of autistic children in LMICs to be empowered, and collaborative partnerships between families and the wider autism community to become established.

Although LMIC contexts pose greater challenges to the well-being and development of families of autistic children, there is limited evidence for caregiver support in these conditions. The promising results in this study indicate the value of research in LMICs using more robust designs, such as randomized controlled trials, to further build the evidence base of family supports for autism outside Western countries, with the ultimate goal of eliminating disparities in population health. Such research should involve larger sample sizes and further examine the treatment effects of programs that are led by trained non-specialists and whether there are differential effects based on levels of autism severity. Future studies should also investigate the moderators and mediators of such programs in order to better identify what works, for whom, under which circumstances. A number of other experimental designs could be used to establish causal links between critical components and outcomes, such as micro-trials (Howe et al., 2010) and factorial experiments within the MOST framework (Collins et al., 2014). Future studies should also document costs and cost-effectiveness to ensure the best use of limited funds. Observational measures and extended follow-up periods are recommended to reduce the source of bias and assess the long-term impact of such programs.

## Conflict of Interest

The authors of this article declare no conflict of interest.

## Acknowledgements

The study would not be possible without the support of the participating families and staff of the Beijing Stars and Rain Education Institute for Autism. We also thank Professor Frances Gardner, Dr Yulia Shenderovich, Professor David Humphreys, and Professor Daniel Michelson for providing insight and expertise that greatly contributed to the design and interpretation of the research. We thank Professor Xiangming Fang, Dr Mark Fransham, Dr Ilan Cerna-Turoff, and Ms Shuangyue Shangguan for their valuable advice on the analytic methods.

## Note

<sup>1</sup>This study was preregistered (<https://clinicaltrials.gov/ct2/show/NCT04257331>). The anonymized individual participant data, analytic code, unpublished full study protocol, and other research materials will be available upon request.

## References

- Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA school-age forms & profiles: Child behavior checklist for ages 6-18, teacher's report form, youth self-report: An integrated system of multi-informant assessment*. Research Center for Children Youth & Families, University of Vermont.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub.
- Anan, R. M., Warner, L. J., McGilivray, J. E., Chong, I. M., & Hines, S. J. (2008). Group intensive family training (GIFT) for preschoolers with autism spectrum disorders. *Behavioral Interventions*, 23(3), 165-180. <https://doi.org/10.1002/bin>
- Appelbaum, M., Cooper, H., Kline, R. B., Mayo-Wilson, E., Nezu, A. M., & Rao, S. M. (2018). Journal article reporting standards for quantitative research in psychology: The APA Publications and Communications Board task force report. *American Psychologist*, 73(1), 3. <https://doi.org/10.1037/AMP0000191>
- Arnold, D. S., O'leary, S. G., Wolff, L. S., & Acker, M. M. (1993). The Parenting Scale: A measure of dysfunctional parenting in discipline situations. *Psychological Assessment*, 5(2), 137. <https://doi.org/10.1037/1040-3590.5.2.137>
- Austin, P. C., & Stuart, E. A. (2015). Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Statistics in Medicine*, 34(28), 3661-3679. <https://doi.org/10.1002/sim.6607>
- Axford, N., Lehtonen, M., Kaoukji, D., Tobin, K., & Berry, V. (2012). Engaging parents in parenting programs: Lessons from research and practice. *Children and Youth Services Review*, 34(10), 2061-2071. <https://doi.org/10.1016/j.childyouth.2012.06.011>
- Bandura, A. (1971). Vicarious and self-reinforcement processes. In R. Glaser (Ed.), *The nature of reinforcement* (pp. 228-278). Academic Press.
- Basten, M., van der Ende, J., Tiemeier, H., Althoff, R. R., Rijlaarsdam, J., Jaddoe, V. W. V., Hofman, A., Hudziak, J. J., Verhulst, F. C., & White, T. (2014). Nonverbal intelligence in young children with dysregulation: The Generation R Study. *European Child and Adolescent Psychiatry*, 23(11), 1061-1070. <https://doi.org/10.1007/s00787-014-0551-x>
- Baxter, A. J., Brugha, T. S., Erskine, H. E., Scheurer, R. W., Vos, T., & Scott, J. G. (2015). The epidemiology and global burden of autism spectrum disorders. *Psychological Medicine*, 45(3), 601-613. <https://doi.org/10.1017/S003329171400172X>
- Bearss, K., Burrell, T. L., Stewart, L., & Scahill, L. (2015). Parent training in autism spectrum disorder: What's in a name? *Clinical Child and Family Psychology Review*, 18(2), 170-182. <https://doi.org/10.1007/s10567-015-0179-5>
- Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition*, 1(1), Article 9. <https://doi.org/10.5334/JOC.10>
- Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implication of propensity score matching. *Journal of Economic Surveys*, 22(1), 31-72. <https://doi.org/10.1111/j.1467-6419.2007.00527.x>
- Carr, T., Shih, W., Lawton, K., Lord, C., King, B., & Kasari, C. (2016). The relationship between treatment attendance, adherence, and outcome

- in a caregiver-mediated intervention for low-resourced families of young children with autism spectrum disorder. *Autism*, 20(6), 643-652. <https://doi.org/10.1177/1362361315598634>
- Chacko, A., Jensen, S. A., Lowry, L. S., Cornwell, M., Chmuklis, A., Chan, E., Lee, D., & Pulgarin, B. (2016). Engagement in behavioral parent training: Review of the literature and implications for practice. *Clinical Child and Family Psychology Review*, 19(3), 204-215. <https://doi.org/10.1007/s10567-016-0205-2>
- Chang, K., & Zaroff, C. M. (2017). Applied behavior analysis in autism spectrum disorders in China and Hong Kong. *Acta Psychopathologica*, 3(5), 0-0. <https://doi.org/10.4172/2469-6676.100124>
- Collins, L. M., Nahum-Shani, I., & Almirall, D. (2014). Optimization of behavioral dynamic treatment regimens based on the sequential, multiple assignment, randomized trial (SMART). *Clinical Trials (London, England)*, 11(4), 426-434. <https://doi.org/10.1177/1740774514536795>
- Cooper, J. O., Heron, T. H., & Heward, W. L. (2007). *Applied behavior analysis* (2nd Edition). Pearson Education Inc.
- De Vries, P. J. (2016). Thinking globally to meet local needs: Autism spectrum disorders in Africa and other low-resource environments. *Current Opinion in Neurology*, 29(2), 130-136. <https://doi.org/10.1097/WCO.0000000000000297>
- Fang, Z., Barlow, J., & Zhang, C. (2022). Parenting programs that address physical abuse in childhood for families of children with developmental disabilities in Mainland China: Systematic review and meta-regression. *Trauma, Violence & Abuse*, 23(2), 457-475. <https://doi.org/10.1177/1524838020915599>
- Fang, Z., Lachman, J. M., Zhang, C., Qiao, D., & Barlow, J. (2022). A virtuous circle: Stakeholder perspectives of a short-term intensive parent training programme delivered within the context of routine services for autism in China. Autism. Advance online publication. <https://doi.org/10.1177/13623613211070869>
- Gardner, F., Montgomery, P., & Knerr, W. (2015). Transporting evidence-based parenting programs for child problem behavior (age 3-10) between countries: Systematic review and meta-analysis. *Journal of Clinical Child & Adolescent Psychology*, 44(1), 1-14. <https://doi.org/10.1080/15374416.2015.1015134>
- Guo, D., Deng, X., Zhao, Q., Wen, H., & Huang, J. (2014). 孤独症家长需求分析与对策建议 [The needs of parents of children with autism in China]. *Disability Research*, 2, 43-48.
- Hartung, J., Knapp, G., & Sinha, B. K. (2008). *Statistical meta-analysis with applications*. Wiley.
- Hastings, R. P., & Brown, T. (2002). Behavior problems of children with autism, parental self-efficacy, and mental health. *American Journal on Mental Retardation*, 107(3), 222-232. [https://doi.org/10.1352/0895-8017\(2002\)107<0222:BPOCWA>2.0.CO;2](https://doi.org/10.1352/0895-8017(2002)107<0222:BPOCWA>2.0.CO;2)
- Howe, G. W., Beach, S. R. H., & Brody, G. H. (2010). Microtrial methods for translating gene-environment dynamics into preventive interventions. *Prevention Science*, 11(4), 343-354. <https://doi.org/10.1007/s11121-010-0177-2>
- Imbens, G. W. (2004). Nonparametric estimation of average treatment effects under exogeneity: A review. *Review of Economics and Statistics*, 86(1), 4-29. <https://doi.org/10.1162/003465304323023651>
- Incredible Years. (2020). Measures and forms for researchers and clinicians | The Incredible Years. <http://www.incredibleyears.com/for-researchers/measures/>
- Leijten, P., Gardner, F., Melendez-Torres, G. J., Van Aar, J., Hutchings, J., Schulz, S., Knerr, W., & Overbeek, G. (2019). Meta-analyses: Key parenting program components for disruptive child behavior. *Journal of the American Academy of Child & Adolescent Psychiatry*, 58(2), 180-190. <https://doi.org/10.1016/j.jaac.2018.07.900>
- Leijten, P., Raaijmakers, M. A. J., Orobio De Castro, B., & Matthys, W. (2013). Does socioeconomic status matter? A meta-analysis on parent training effectiveness for disruptive child behavior. *Journal of Clinical Child & Adolescent Psychology*, 42(3), 384-392. <https://doi.org/10.1080/15374416.2013.769169>
- Li, F., & Thomas, L. E. (2019). Addressing extreme propensity scores via the overlap weights. *American Journal of Epidemiology*, 188(1), 250-257. <https://doi.org/10.1093/aje/kwy201>
- Liao, Y., Dillenburg, K., He, W., Xu, Y., & Cai, H. (2020). A systematic review of applied behavior analytic interventions for children with autism in mainland China. *Review Journal of Autism and Developmental Disorders*, 7(4), 333-351. <https://doi.org/10.1007/S40489-020-00196-W/TABLES/3>
- Liao, Y., Dillenburg, K., & Hu, X. (2022). Behavior analytic interventions for children with autism: Policy and practice in the United Kingdom and China. *Autism*, 26(1), 101-120. <https://doi.org/10.1177/13623613211020976>
- Lin, L., & Aloe, A. M. (2021). Evaluation of various estimators for standardized mean difference in meta-analysis. *Statistics in Medicine*, 40(2), 403-426. <https://doi.org/10.1002/SIM.8781>
- Liu, J., Cheng, H., & Leung, P. W. L. (2011). The application of the preschool child behavior checklist and the Caregiver-Teacher Report Form to Mainland Chinese children: Syndrome structure, gender differences, country effects, and inter-informant agreement. *Journal of Abnormal Child Psychology*, 39(2), 251-264. <https://doi.org/10.1007/s10802-010-9452-8>
- Liu, P., & Liu, J. (2018). 自闭症群体的家庭需求与支持体系构建 [Needs in families of children with Autism and the establishment of support system]. *Academic Exchange*, 8, 113-121.
- Lovibond, S. H., & Lovibond, P. F. (1995). *Manual for the Depression Anxiety & Stress Scales* (2nd ed.). Psychology Foundation.
- Maenner, M. J., Shaw, K. A., Baio, J., Washington, A., Patrick, M., DiRienzo, M., Christensen, D. L., Wiggins, L. D., Pettygrove, S., Andrews, J. G., Lopez, M., Hudson, A., Baroud, T., Schwenk, Y., White, T., Rosenberg, C. R., Lee, L.-C., Harrington, R. A., Huston, M., ... Dietz, P. M. (2020). Prevalence of autism spectrum disorder among children aged 8 years. *MMWR. Surveillance Summaries*, 69(4), 1-12. <https://doi.org/10.15585/mmwr.ss6904a1>
- McCabe, H. (2012). Bamboo shoots after the rain: Development and challenges of autism intervention in China. *Autism*, 17(5), 510-526. <https://doi.org/10.1177/1362361312436849>
- McEachern, A. D., Dishion, T. J., Weaver, C. M., Shaw, D. S., Wilson, M. N., & Gardner, F. (2012). Parenting Young Children (PARYC): Validation of a self-report parenting measure. *Journal of Child and Family Studies*, 21(3), 498-511. <https://doi.org/10.1007/s10826-011-9503-y>
- McIntyre, L. L. (2008). Adapting Webster-Stratton's Incredible Years parent training for children with developmental delay: Findings from a treatment group only study. *Journal of Intellectual Disability Research*, 52(12), 1176-1192. <https://doi.org/10.1111/j.1365-2788.2008.01108.x>
- Mesibov, G. B., & Shea, V. (2010). The TEACCH Program in the era of evidence-based practice. *Journal of Autism and Developmental Disorders*, 40(5), 570-579. <https://doi.org/10.1007/s10803-009-0901-6>
- Mockford, C., & Barlow, J. (2004). Parenting programmes: Some unintended consequences. *Primary Health Care Research and Development*, 5(3), 219-227. <https://doi.org/10.1191/1463423604pc2000a>
- Morris, J., Belfer, M., Daniels, A., Flisher, A., Villé, L., Lora, A., & Saxena, S. (2011). Treated prevalence of and mental health services received by children and adolescents in 42 low-and-middle-income countries. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 52(12), 1239-1246. <https://doi.org/10.1111/j.1469-7610.2011.02409.x>
- National Bureau of Statistics of China. (2011). *Communiqué of the National Bureau of Statistics of People's Republic of China on major figures of the 2010 population census* (No. 1).
- National Research Council. (2001). *Educating children with autism*. Committee on Educational Interventions for Children with Autism. Division of Behavioral and Social Sciences and Education (Ed.). National Academy Press.
- Nevill, R. E., Lecavalier, L., & Stratis, E. A. (2018). Meta-analysis of parent-mediated interventions for young children with autism spectrum disorder. *Autism*, 22(2), 84-98. <https://doi.org/10.1177/1362361316677838>
- Oakley, A., Strange, V., Bonell, C., Allen, E., & Stephenson, J. (2006). Process evaluation in randomised controlled trials of complex interventions. *British Medical Journal*, 332(7538), 413-416. <https://doi.org/10.1136/bmj.332.7538.413>
- Oono, I. P., Honey, E. J., & McConachie, H. (2013). Parent-mediated early intervention for young children with autism spectrum disorders (ASD). *Evidence-Based Child Health: A Cochrane Review Journal*, 8(6), 2380-2479. <https://doi.org/10.1002/14651858.CD009774.pub2.www.cochranelibrary.com>
- Pandolfi, V., Magyar, C. I., & Dill, C. A. (2009). Confirmatory factor analysis of the Child Behavior Checklist 1.5-5 in a sample of children with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 39(7), 986-995. <https://doi.org/10.1007/s10803-009-0716-5>
- Postorino, V., Sharp, W. G., McCracken, C. E., Bearss, K., Burrell, T. L., Evans, A. N., & Scahill, L. (2017). A systematic review and meta-analysis of parent training for disruptive behavior in children with autism spectrum disorder. *Clinical Child and Family Psychology Review*, 20(4), 391-402. <https://doi.org/10.1007/s10567-017-0237-2>
- Quinn, M., Carr, A., Carroll, L., & O'Sullivan, D. (2007). Parents Plus Programme 1: Evaluation of its effectiveness for pre-school children with developmental disabilities and behavioural problems. *Journal of Applied Research in Intellectual Disabilities*, 20(4), 345-359. <https://doi.org/10.1111/j.1468-3148.2006.00352.x>
- Reichow, B., Hume, K., Barton, E. E., & Boyd, B. A. (2018). Early intensive behavioral intervention (EIBI) for young children with autism spectrum disorders (ASD). *Cochrane Database of Systematic Reviews*, 5. <https://doi.org/10.1002/14651858.CD009260.pub3>
- Rimland, B., & Edelson, S. (1999). *Autism Treatment Evaluation Checklist (ATEC)*. Autism Research Institute.
- Rubin, D. B. (2001). Using propensity scores to help design observational studies: Application to the tobacco litigation. *Health Services and Outcomes Research Methodology*, 2(3-4), 169-188. <https://doi.org/10.1023/A:1020363010465>
- Schreibman, L., Dawson, G., Stahmer, A. C., Landa, R., Rogers, S. J., McGee, G. G., Kasari, C., Ingersoll, B., Kaiser, A. P., Bruinsma, Y., McNerney, E., Wetherby, A., & Halladay, A. (2015). Naturalistic developmental behavioral interventions: Empirically validated treatments for autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 45(8), 2411-2428. <https://doi.org/10.1007/s10803-015-2407-8>
- Simonoff, E., Pickles, A., Charman, T., Chandler, S., Loucas, T., & Baird, G. (2008). Psychiatric disorders in children with autism spectrum

- disorders: Prevalence, comorbidity, and associated factors in a population-derived sample. *Journal of the American Academy of Child and Adolescent Psychiatry*, 47(8), 921-929. <https://doi.org/10.1097/CHI.0b013e318179964f>
- Skinner, B. F. (1950). Are theories of learning necessary? *Psychological Review*, 57(4), 193-216. <https://doi.org/10.1037/h0054367>
- Skotarczak, L., & Lee, G. K. (2015). Effects of parent management training programs on disruptive behavior for children with a developmental disability: A meta-analysis. *Research in Developmental Disabilities*, 38, 272-287. <https://doi.org/10.1016/j.ridd.2014.12.004>
- Sloven, N., Fitzmaurice, G., Williams, D. R., & Gilman, S. E. (2010). Poverty, food insecurity, and the behavior for childhood internalizing and externalizing disorders. *Journal of the American Academy of Child & Adolescent Psychiatry*, 49(5), 444-452. <https://doi.org/10.1016/j.jaac.2010.01.018>
- Smilkstein, G., Ashworth, C., & Montano, D. (1982). Validity and reliability of the family APGAR as a test of family function. *Journal of Family Practice*, 15(2), 303-311.
- Strauss, K., Mancini, F., SPC Group, & Fava, L. (2013). Parent inclusion in early intensive behavior interventions for young children with ASD: A synthesis of meta-analyses from 2009 to 2011. *Research in Developmental Disabilities*, 34(9), 2967-2985. <https://doi.org/10.1016/j.ridd.2013.06.007>
- Stuart, E. A., Huskamp, H. A., Duckworth, K., Simmons, J., Song, Z., Chernew, M. E., & Barry, C. L. (2014). Using propensity scores in difference-in-differences models to estimate the effects of a policy change. *Health Services and Outcomes Research Methodology*, 14(4), 166-182. <https://doi.org/10.1007/s10742-014-0123-z>
- Tellegen, C. L., & Sanders, M. R. (2013). Stepping Stones Triple P-Positive Parenting Program for children with disability: A systematic review and meta-analysis. *Research in Developmental Disabilities*, 34(5), 1556-1571. <https://doi.org/10.1016/j.ridd.2013.01.022>
- Totsika, V., Hastings, R. P., Emerson, E., Lancaster, G. A., & Berridge, D. M. (2011). A population-based investigation of behavioural and emotional problems and maternal mental health: Associations with autism spectrum disorder and intellectual disability. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 52(1), 91-99. <https://doi.org/10.1111/j.1469-7610.2010.02295.x>
- van Aar, J., Leijten, P., de Castro, O. B., & Overbeek, G. (2017). Sustained, fade-out or sleeper effects? A systematic review and meta-analysis of parenting interventions for disruptive child behavior. *Clinical Psychology Review*, 51, 153-163. <https://doi.org/10.1016/j.cpr.2016.11.006>
- Williams, M. E., Hastings, R. P., & Hutchings, J. (2020). The Incredible Years autism spectrum and Language Delays Parent Program: A pragmatic, feasibility randomized controlled trial. *Autism Research*, 13(6), 1011-1022. <https://doi.org/10.1002/aur.2265>
- Wing, C., Simon, K., & Bello-Gomez, R. A. (2018). Designing difference in difference studies: Best practices for public health policy research. *Annual Review of Public Health*, 39. <https://doi.org/10.1146/annurev-publhealth-040617-013507>
- World Health Organization. (2020). *Process of translation and adaptation of instruments*. [https://www.who.int/substance\\_abuse/research\\_tools/translation/en/](https://www.who.int/substance_abuse/research_tools/translation/en/)
- Xiong, N., Ji, C., Bo, H., Wang, X., Shi, J., Chen, Y., & Li, Y. (2010). 孤独症儿童临床症状及母亲心理状况的应用行为分析干预效果 [The effectiveness of applied behavior analysis intervention on clinic symptoms of children with autism and psychological condition of their mothers]. *Chinese Journal of Rehabilitation Medicine*, 25(10), 564-566.
- Zhou, H., Xu, X., Yan, W., Zou, X., Wu, L., Luo, X., Li, T., Huang, Y., Guan, H., Chen, X., Mao, M., Xia, K., Zhang, L., Li, E., Ge, X., Zhang, L., Li, C., Zhang, X., Zhou, Y., ... Wang, Y. (2020). Prevalence of autism spectrum disorder in China: A nationwide-multi-center population-based study among children aged 6 to 12 years. *Neuroscience Bulletin*, 36(9), 961-971. <https://doi.org/10.1007/S12264-020-00530-6/TABLES/4>