ORIGINAL PAPER

Randomized Controlled Trial of the Focus Parent Training for Toddlers with Autism: 1-Year Outcome

Iris Oosterling · Janne Visser · Sophie Swinkels ·
Nanda Rommelse · Rogier Donders · Tim Woudenberg ·
Sascha Roos · Rutger Jan van der Gaag · Jan Buitelaar

Abstract This randomized controlled trial compared results obtained after 12 months of nonintensive parent training plus care-as-usual and care-as-usual alone. The training focused on stimulating joint attention and language skills and was based on the intervention described by Drew et al. (Eur Child Adolesc Psychiatr 11:266–272, 2002). Seventy-five toddlers with autism spectrum disorder (65 autism, 10 PDD-NOS, mean age = 34.4 months, SD = 6.2) were enrolled. Analyses were conducted on a final sample of 67 children (lost to follow-up = 8). No significant intervention effects were found for any of the primary (language), secondary (global clinical improvement), or mediating (child engagement, early precursors of social communication, or parental skills) outcome variables, suggesting that the 'Focus parent training' was not of additional value to the more general care-as-usual.

Keywords Autism · Parent training · Toddler · Early intervention · Randomized controlled trial

I. Oosterling (✉) · J. Visser · S. Swinkels · N. Rommelse ·
T. Woudenberg · S. Roos · R. J. van der Gaag · J. Buitelaar
Karakter Child and Adolescent Psychiatry University Centre,
Reinier Postlaan 12, 6525 GC Nijmegen, The Netherlands
e-mail: i.oosterling@karakter.com

S. Swinkels · N. Rommelse · R. J. van der Gaag
Department of Psychiatry, Nijmegen Medical Centre, Radboud
University, Nijmegen, The Netherlands

S. Swinkels · N. Rommelse · R. J. van der Gaag · J. Buitelaar
Nijmegen Centre for Evidence Based Practice (NCEBP),
Nijmegen, The Netherlands

R. Donders
Department of Epidemiology, Biostatistics and HTA, Nijmegen
Medical Centre, Radboud University, Nijmegen, The
Netherlands

R. J. van der Gaag · J. Buitelaar
Department of Cognitive Neuroscience, Nijmegen Medical
Centre, Radboud University, Nijmegen, The Netherlands

Present Address:
S. Swinkels
Danone Research, Centre for Specialised Nutrition,
Wageningen, The Netherlands

Introduction

In clinical practice there is consensus that children with autism spectrum disorder (ASD) should be enrolled into intervention programs as early as possible (Dawson and Osterling 1997; National Research Counsil 2001; Rogers and Vismara 2008); however, there is little empirical evidence that earlier intervention is more beneficial than later intervention (Charman 2003). The need for empirically based early intervention programs has become more urgent because substantial advances have been made in the early detection and diagnosis of ASD (Charman and Baird 2002; Oosterling et al. 2010).

Core deficits in the areas of social interaction, and language and communication are already present in infants and toddlers with ASD. Several longitudinal studies have shown early social communication skills to be associated with language outcomes (e.g., Bono et al. 2004, Charman et al. 2003; Dawson et al. 2004; Mundy et al. 1990; Toth et al. 2006), making these skills feasible targets for intervention. For this reason, many early intervention programs focus on promoting skills, such as joint attention, imitation, and play, which are considered prerequisites for the development of social communicative abilities in both typically developing children and children with autism.

According to the National Research Counsil (2001), it is essential to include parent training in early intervention programs because parents generally represent the most proximal and powerful environmental influence during early childhood, and without parental participation gains in children's development are unlikely to be maintained (Bruner 1981; Whalen et al. 2006). McConachie and Diggle (2007) systematically reviewed the evidence of parent-implemented interventions for children aged 1–6 years with ASD. Results from (randomized) controlled studies (n = 10) demonstrated that parent training could improve child communication behavior (Aldred et al. 2004; Drew et al. 2002), increase maternal knowledge of autism (Jocelyn et al. 1998), reduce maternal depression (Bristol et al. 1993), enhance parent–child interaction, and improve maternal communication style with their child (Aldred et al. 2004; Koegel et al. 1996). The latter has also been described as 'parental sensitivity', which has been suggested to promote communication skills in children with autism (Siller and Sigman 2002). However, the majority of the studies reviewed suffered from methodological shortcomings, such as small sample sizes, absence of long-term follow-up assessments of at least 1 year, limited use of tools with established validity with an ASD diagnosis such as the Autism Diagnostic Observation Schedule (ADOS; Lord et al. 2001) and the Autism Diagnostic Interview-Revised (ADI-R; Rutter et al. 2003), and an inability to explore mediating or moderating effects. Further, some studies did not use a randomized design, which precludes drawing firm conclusions about the effectiveness of the interventions investigated.

The parent training program evaluated for effectiveness in the current study, called *Focus parent training*, is modeled after an intervention described by Drew et al. (2002). This home-based parent-training program focuses on promoting compliance, mutual enjoyment, joint attention (as early precursors of social and communicative behavior), and language development. The program adopted a consultant model with parents acting as the everyday therapist. In the study by Drew et al. parents were randomized to the experimental parent-training group (n = 12) or to local services only (n = 12). After 12 months enrolment in the study (mean age of the children = 35 months), results suggested that children in the parent training group made more progress in language development than the children in the local services group. No differences were found for symptom severity or parent self-reported stress. Limitations of this study are the small sample size, which may have resulted in findings by chance, and outcome data being based on parental report only, which can introduce correlated measurement error (parents have been trained in only one group, and these trained parents can become more sensitive to evidence of their children's understanding and word

approximations regardless of their children's real development). Therefore, the promising findings warrant replication.

The present study sought to avoid the methodological shortcomings of previous studies by using a number of widely used professional- and parent-report measures with acceptable validity and a larger sample (N = 67). Based on an alpha of 0.05 and a beta of 0.80, a sample size of 34 participants was required to detect a treatment effect of medium size within one group. A sample of 68 participants was required to detect a difference between the groups. The 2-year program was evaluated 1 year after the start of the intervention. The aim was to replicate and extend the findings of Drew et al. (2002), including subgroup analyses based on degree of developmental delay at baseline in order to establish whether some children benefited more than others, based on their developmental potential. Based on the assumption that joint attention is a predictor of concurrent language ability (Dawson et al. 2004), we hypothesized that the Focus parent training would have a specific, positive effect on the development of language.

Methods

Participants

Seventy-five children with ASD were randomized to the experimental or control group, but 8 were lost to endpoint measurements (see Fig. 1). Participants were recruited between spring 2004 and spring 2007 at Karakter Child and Adolescent Psychiatry University Center Nijmegen, the Netherlands. They were referred by clinicians because of possible ASD, as identified by screen positive results on the Early Screening of Autistic Traits Questionnaire (ESAT; Oosterling et al. 2009; Swinkels et al. 2006). The ESAT is a 14-item screening instrument to identify very young children at risk for ASD. Consensus diagnosis was made by at least two experienced, board-certified professionals (a child psychiatrist and a psychologist), subsequent to a comprehensive diagnostic assessment procedure including clinical observations of the child using the ADOS (Lord et al. 2001), the parent interview ADI-R (Rutter et al. 2003), and psychometric testing of developmental abilities. These were measured with the *Mullen Scales of Early Learning* (MSEL; Mullen 1995) in 22 children (29%). The remaining children, who were difficult to test with the MSEL were assessed with the *Psycho Educational Profile-Revised (PEP-R*; Schopler et al. 1990). The MSEL is a developmental test with adequate reliability and validity intended for use in children aged 0–68 months who are willing to cooperate and yields an Early Learning Composite score (mean = 100, SD = 15). The PEP-R was used for children with very limited receptive and expressive
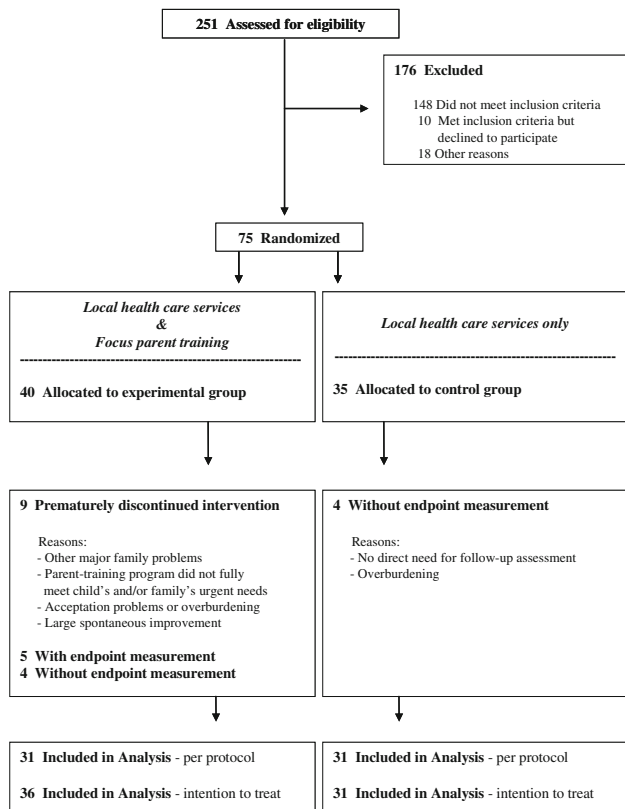
Fig. 1 Flow diagram of participants through each stage of the trial

language and who could only be engaged indirectly by their reactions to the test materials. The PEP-R offers a developmental approach yielding a profile of well established skills and emerging skills for children aged 6 months–7 years. Developmental Quotient (DQ) based on the PEP-R was calculated as: (developmental age in months/chronological age in months)*100. Interrater reliability of the PEP-R is high, and with regard to validity significant correlations were found with other developmental tests such as the Bayley Scales of Infant Development ($r = 0.77$; Bayley 1969).

Inclusion criteria for the Focus parent training were chronological age (12–42 months), clinical diagnosis, and developmental level. The latter criterion was used because the parent training is not appropriate for children with a developmental age younger than 12 months or with extremely little developmental potential, or for children with limited room for improvement concerning target behaviors at the upper end of the spectrum. Therefore, we only included children with either a diagnosis of autism in combination with a developmental age of at least 12 months or children with a diagnosis of Pervasive Developmental Disorder-Not Otherwise Specified (PDD-NOS) in combination with a developmental age of at least 12 months and a Developmental Quotient (DQ) below 80.

Exclusion criteria were substantial problems within the family, other than those related to the child's condition (such as severe parental psychopathology, financial/housing problems, and marital conflicts), which could interfere with the parent training, and insufficient parental proficiency in Dutch.

Eight participants did not fully meet the inclusion criteria but were included anyway because it was felt that there was enough room for improvement based on clinical assessment, either at the lower or upper limits: 4 children in the experimental group and 3 in the control group had autism but had a developmental age younger than 12 months. In addition, 1 child included in the experimental group was diagnosed with PDD-NOS but had a DQ of 86 at baseline. Of these 8 participants 2 had no endpoint measurement.

Randomization and Baseline Characteristics

Participants randomized to the experimental group received the Focus parent training in addition to care-as-usual, and those randomized to the control group received care-as-usual alone. Participants were randomized after completion of baseline assessments and after signing for informed consent, using two strategies. The first 26 participants were randomly assigned to one of the groups, but we learned that the children and parents in the two groups could come into contact with each other (the children attended the same specialized day-care nurseries) and might exchange experiences, thereby potentially undermining the distinction between the experimental and control conditions. For this reason, we allocated the remaining participants ($n = 49$) by where participants lived. Subjects living outside the Nijmegen area were allocated to the control group, while subjects living inside that area were allocated to the experimental group. Clusters (living inside or outside the regions specified) were not randomized for pragmatic reasons (i.e., travel distance for parent trainers). At baseline, minimal differences regarding clinical features of the participants were found between groups. See Table 1 for more details.

Care-as-Usual

In the Netherlands, daycare for young children with developmental problems is provided through either (a) *special daycare centers* for children with mental retardation (in combination, or not, with other developmental, psychiatric, or medical conditions) or (b) *medical nurseries* for children with behavioral or developmental problems, but without developmental delay. Both settings provide, on an individual basis, speech and language therapy, motor therapy, music therapy, and play therapy. Support for

Table 1 Participant characteristics at baseline and care-as-usual received from baseline to endpoint

	Experiment group (n = 36)		Control group (n = 31)		t	χ^2
	Mean (SD)	(%)	Mean (SD)	(%)		
Child characteristics						
Age in months	35.2 (5.5)		33.3 (6.4)		−1.31	
Male		75.0		80.6		0.31
Diagnosis						
Autism		91.7		83.9		0.96
PDD-NOS		8.3		16.1		
DQ[a]	58.4 (16.8)		58.0 (16.9)		−0.11	
ADI-R[a]						
RSI	16.3 (5.1)		14.7(4.5)		−1.27	
Comm	11.2 (2.4)		10.3 (2.8)		−1.45	
RRSPB	**4.1 (2.1)**		**3.0 (1.8)**		**−2.31***	
ADOS[a]						
SA	15.0 (4.6)		14.8 (4.9)		−0.14	
RRB	2.8 (1.7)		2.8 (1.9)		−0.11	
MacArthur N-CDI						
Words understood	177.9 (122.5)		181.5 (121.4)		0.12	
Words said	106.8 (122.2)		101.7 (109.7)		−0.18	
Gestures produced	29.1 (13.7)		30.1 (13.6)		0.29	
Erikson Scales[b]						
Non-negativity	5.9 (1.8)		6.2 (0.8)		0.87	
Non-avoidance	3.9 (1.5)		4.1 (1.3)		0.38	
Compliance	3.8 (1.6)		4.2 (1.3)		0.89	
CBCL[c]						
Internalizing	**21.3 (9.4)**		**16.9 (7.3)**		**−2.05***	
Externalizing	21.2 (11.1)		19.4 (9.0)		−0.71	
ICQ[d]						
Total score	146.4 (27.0)		141.0 (18.0)		−0.82	
Family/parent characteristics						
Non-western immigrants		91.7		83.9		0.96
Single parents families		91.7		96.8		0.77
Mothers' educational level						
Low		41.7		41.9		0.06
Middle		33.3		35.5		
High		25.0		22.6		
Fathers' educational level[a]						
Low		**34.3**		**56.7**		**6.34***
Middle		**20.0**		**26.7**		
High		**45.7**		**16.7**		
Erikson Scales[b]						
Supportive presence	4.5 (1.3)		4.8 (1.3)		0.81	
Respect for authonomy	4.9 (1.0)		4.8 (1.4)		−0.42	
Structure and limit setting	4.2 (1.3)		4.2 (1.4)		0.08	
Quality of instruction	4.3 (1.4)		4.3 (1.3)		0.06	
Non-hostility	6.7 (0.7)		6.9 (0.3)		1.68	
SCL-90						
Mothers (n = 57)	126.7 (31.2)		123 (28.0)		−0.46	
Fathers (n = 47)	113.2 (33.7)		112.3 (21.9)		−0.10	

Table 1 continued

	Experiment group ($n = 36$)		Control group ($n = 31$)		t	χ^2
	Mean (SD)	(%)	Mean (SD)	(%)		
Stress index total (NOSI)						
Mothers ($n = 58$)	343.7 (84.2)		340.0 (84.1)		−0.17	
Fathers ($n = 46$)	307.2 (94.4)		333.9 (82.9)		0.97	
Care-as-usual						
Day care[e]	5.2 (1.7)		4.2 (2.9)		−1.68	
Speech and language th.[f]	16.7 (22.4)		19.1 (22.0)		0.43	
Physical therapy[f]	8.3 (18.4)		6.4 (14.9)		−0.46	
Other individual therapy[f]	24.9 (59.5)		22.7 (39.7)		−0.17	
Parental counseling[f]	21.0 (30.9)		28.2 (36.2)		0.88	

Note. PDD-NOS = pervasive developmental disorder—not otherwise specified; DQ = Developmental Quotient; ADI-R = Autism Diagnostic Interview-Revised (Rutter et al. 2003); RSI = Reciprocal Social Interaction total; Comm = Communication total; RRSPB = Restricted, Repetitive, and Stereotyped Patterns of Behavior total; ADOS = Autism Diagnostic Observation Schedule (Lord et al. 2001); SA = Social Affect total (revised algorithm, Gotham et al. 2007); RRB = Restricted, Repetitive Behavior total (revised algorithm); N-CDI = Dutch version of the MacArthur Communicative Development Inventory (Fenson et al. 1993; Zink and Lejaegere 2002); CBCL = Child Behavior Checklist 1½−5 (Achenbach and Rescorla 2000); SCL-90 = Symptom Checklist-90 (Arrindell and Ettema 1975); NOSI = Nijmeegse Ouderlijke Stress Index (de Brock et al. 1992); ICQ = Infant Characteristics Questionnaire (Bates et al. 1979)

* $p < 0.05$ (printed in bold)

[a] No value = 2

[b] No value = 6, and scores are only displayed for mother–child interactions. For father-child interactions no significant differences between groups were found either

[c] No value = 4

[d] No value = 10

[e] Average number of daily periods (morning/afternoon) spend in either a child special day care centre or a medical nursery

[f] Minutes per week

parents can range from low-frequency sessions with a psychologist (e.g., 1 h per month) to intensive practical support set up in the home environment (e.g., twice a week for 90 min). There were no significant differences in the care-as-usual received between the control and experimental groups (see Table 1).

Aims and Content of the Focus Parent Training

General Approach and Aims

The Focus parent training is a 2-year intervention program using a professional-as-consultant and parent-as-therapist model and adopting an eclectic approach within a social-pragmatic and developmental context (Ospina et al. 2008). Auriol Drew (former speech and language therapist at Guy's Hospital, Newcomen Centre, London) provided in-company training for the parent-trainers prior to the start of the study. In total eight psychologists or sociotherapists worked as parent-trainers. Regular meetings were held so that the parent-trainers could discuss difficulties encountered during the interventions.

The aims of the parent training at a *child level* were threefold: to promote the child's engagement (compliance and willingness to join in mutual activities), to elicit early precursors of social communication (joint pleasure and joint attention behaviors, imitation, and functional play), and to stimulate language development. At a *parent level*, the aim of the training was to stimulate parental skills in order to promote child development (see Intervention Techniques taught to Parents).

Training Scheme

The parent-training program started with four weekly 2-h sessions with a group of parents, followed by individual 3-h home visits every 6 weeks during the first year. In the second year, the home visits were scheduled at 3-month intervals. Plenary sessions were held every 6 months, during which the parent-trainers presented the principles of behavioral management and the social pragmatic approach taken to the development of joint attention, nonverbal social communication, and language skills. These sessions also gave parents the opportunity to share emotions and experiences.

Home visits followed a set structure including (a) parent–child free play, followed by discussion and feedback (often based on video material), (b) practicing specific games (see

below), and (c) discussion of possibilities to facilitate generalization into daily routines, e.g., during mealtimes, dressing, or shopping. Each home visit was concluded by setting short-term goals and choosing corresponding activities to practice during the next 6-week period.

Intervention Activities

Parents were encouraged to keep the child engaged in mutual activities, either during free play (child led), during specific game play (parent led), or in everyday joint action routines. Specific games, all of which had written instructions, included (a) games to illicit gaze switching and eye contact, e.g., with balloons or bubbles; (b) games to illicit index finger pointing and showing; (c) games to stimulate first words (e.g., picture posting and sorting, lotto, puzzles); (d) mirror games to stimulate dyadic joint attention, imitation, or play with objects; (e) games for learning to greet others, and; (f) games for learning to be curious and using words to show curiosity, such as 'what', 'who', and 'where'. All specific games were designed in such a way that they could be used for different developmental levels.

Initially, 2–3 min were recommended for each activity, gradually increasing to about 5 min per activity. Parents were advised to take between 30 and 60 min "set-aside" time per day. However, it was anticipated that as parents became more familiar with the activities and techniques, this social pragmatic approach would be used increasingly often and for longer during everyday parent–child interactions.

Intervention Techniques Taught to Parents

The parents of the children in the intervention group were taught several techniques to maximize the effect of the intervention.

1. Behavior management techniques, including principles of (differential) reinforcement, interrupting unwanted behavior, and teaching alternative behaviors.
2. Techniques to stimulate mutual enjoyment and eye contact including using an (overly) enthusiastic voice, holding objects close to own eyes, and then wait for eye contact before complying with the child's requests.
3. Holistic learning of language, for example use of exaggerated prosody and repetitive paraphrasing to maximize the likelihood that the child would understand the meaning of key referent words.
4. Use of visual support for spoken language (objects, photos, pictures) and use of simple gestures.
5. Consequent attention to adequate pace, timing, and adjustment to the child's developmental level and interests (sensitivity).

Measures

We used a multi-informant perspective on data collection by using professional observation, parent report, and video recording, applying widely used and standardized measures, most of them with acceptable validity. Table 2 provides a summary of measures and information concerning blinding.

Primary Outcome Measures

Language Development Based on parent report, changes in language comprehension and language production were assessed with a Dutch version of the *MacArthur Communicative Development Inventory*, the *N-CDI* (Fenson et al. 1993; Zink and Lejaegere 2002). The N-CDI has adequate reliability (internal consistency) and good criterion validity

Table 2 Summary of measures and blinding

Measures	Time		Approach	Informant	Blind to group status at endpoint?	
	BL	EP			Administation	Coding
Primary outcome measures						
MacArthur N-CDI language	×	×	Parent report	Parent(s)	No	–
ADOS (item A1)	×	×	Clinical observation	Psychologist	No	Yes (89%)
Secondary outcome measure						
CGI-I	–	×	Rating scale	Psychologist	–	Yes (100%)
Mediating outcome measures						
Erickson child scales	×	×	Video observation	Student/junior psychologist	Yes (100%)	Yes (100%)
MacArthur N-CDI gestures	×	×	Parent report	Parent(s)	No	–
ADOS (subscales)	×	×	Clinical observation	Psychologist	No	Yes (89%)
Erickson parent scales	×	×	Video observation	Student/junior psychologist	Yes (100%)	Yes (100%)

Note. BL = baseline; EP = endpoint; ADOS = Autism Diagnostic Observation Schedule (Lord et al. 2001); N-CDI = Dutch version of the MacArthur Communicative Development Inventory (Fenson et al. 1993; Zink and Lejaegere 2002); CGI-I = Clinical Global Impression-Improvement scale (Guy 1976)

as evaluated against a language measure that uses professional observation, the Dutch Non-Speech Test (NNST; Zink and Lembrecht 2000). Based on professional observation, change in language production was assessed with ADOS item A1 (level of non-echoed language). To create unity in scores for modules 1 and 2, we recoded item scores into a 6-point rating scale ranging from score 0 (no words or word approximations) to score 5 (phrase speech of 3 or more words).

Secondary Outcome Measure

General Improvement To assess the degree of general improvement relative to the baseline state, the *Clinical Global Impression—Improvement scale* (*CGI-I*; Guy 1976) was rated by a psychologist blind to the case–control status and based on comprehensive assessment reports. The CGI-I is a clinical outcome measure with acceptable validity that is sensitive to change (Berk et al. 2008). The scale contains a 7-point rating with responses from 1 (very much improved) through to 7 (very much worse).

Mediating Outcome Measures

Engagement Change in child compliance and willingness to join in mutual activities was measured using three 7-point rating scales developed by Erickson et al. (1985): (a) *Compliance* or the child's tendency to follow directions and comply to the requests given by the parent. At the high end, a child complies in a detailed fashion to the directions given by the parent, and at the low end the child actively rejects almost all directions given by the parent; (b) *Negativity* or the child's anger, dislike or hostility in the interaction with the parent. We used reversed scores (Non-negativity). At the high end the child shows neither overt nor covert signs of negativism, and at the low end the child is repeatedly and overtly anger or resistant towards the parent; and (c) *Avoidance* or the child's tendencies or clear attempts in the session to avoid interacting with the parent. We used reversed scores (Non-avoidance). At the high end the child shows no withdrawal from, or intention to avoid the parent, and at the low end the child shows strong interest to withdraw from the parent, either by leaving the situation or resisting the parents' attempts to engage him or her. Scores were based on videotaped parent–child interaction episodes of about 15 min, either at the clinic (baseline measurement) or at home (endpoint measurement), with a standardized set of play material. The videotaped interactions were rated by alternate combinations of pairs of trained students or junior psychologists (5 raters in total) who were blind to the case–control status. The inter-rater agreement (individual coding evaluated against a consensus coding as established by alternate combinations of pairs), based on 42% of the data

and expressed as weighed Cohen's Kappa's (agreement within one scale point), ranged from 57.4 to 74.9 on these three subscales. The clinical significance of Kappa is interpreted as: <40 = poor, 40–59 = fair, 60–74 = good, ≥75 = excellent (Cicchetti and Sparrow 1981).

Early Precursors of Social Communication Change in early precursors of social communication was identified with two measures: (a) The *NCD-I Gestures* form (Fenson et al. 1993), which asks parents to record the communicative and symbolic gestures the child has tried or completed (adequate internal consistency and criterion validity); and (b) the *ADOS* (Lord et al. 2001). At baseline and endpoint, children were assessed with the most appropriate module, depending on their language level. Consequently, there could be differences per child in the module used over time (either Module 1 or Module 2). For this reason, the revised algorithms proposed by Gotham et al. (2007) were applied, which allow for inter-module comparisons of domain scores. In the current study, we used separate items comprising the Social Affect domain score and a sum score of items that comprise a joint attention factor (for more details Gotham et al. 2007). The ADOS was administered by trained psychologists not blind to the case–control status of participants, but was coded based on videotapes by psychologists who met standard requirements for research reliability and who were blind to the case–control status. The ADOS has adequate internal consistency and inter rater reliability, and high sensitivity and specificity regarding diagnostic discrimination (Lord et al. 2001).

Parental Skills Change in the quality of parental skills (five dimensions) in their interaction with their child was examined using the 7-point *Erickson Scales* (Erickson et al. 1985). To this end, videotaped parent–child interaction episodes of about 15 min (see above) were rated by alternate pairs of trained students or junior psychologists blind to the case–control status of the participants. Inter-rater agreement was established in the same way as for the engagement measures. The five parental dimensions were (a) *supportive presence* or the provision of positive regard and emotional support ($\kappa = 80.2$); (b) *respect for the child's autonomy* or de degree to which the parent acts in a way that recognizes and respects the validity of the child's individuality, motives, and perspectives in the session ($\kappa = 55.3$); (c) *effective structure and limit setting* or the degree to which the parent attempts to establish his/her expectations for the child's behavior versus not communicating his/her expectations or not enforcing his/her agenda adequately ($\kappa = 82.6$); (d) *quality of instructions* or the degree to which the parent structures the situation so that the child knows what the task objectives are and receives hints or corrections while solving the problems

that are well timed and paced, graded in logical steps, and stated clearly ($\kappa = 70.1$); and (e) *Non-hostility* that reflects the parent's expression of anger, discounting or rejecting of the child ($\kappa = 100$).

Statistical Analyses

Repeated-measures ANOVAs were used to determine change in ratings before and after treatment, and between the experimental group and the control group for the *primary outcome measure* (language) and for the *mediating outcome measures* (engagement, early precursors of social communication, and parental skills). In these models, group (experimental versus control) was entered as between-subject factor, time (baseline versus endpoint) as the within-subject factor, and the group by time interaction represented the treatment effect. In addition, Developmental Quotient (DQ) divided into three groups (<50; $50–69$; ≥ 70) was included in the model as covariate. The DQ by group by time interaction indicated whether DQ had a moderating effect on the main treatment outcome. With regard to the *secondary outcome measure* (clinical global improvement), we performed a Chi-square analysis to determine whether there was any difference in general improvement between the experimental and control groups relative to the baseline state. Level of significance was set at $p < 0.05$. Analyses were run both on the basis of the 'intention to treat principle' and 'per protocol'. In addition, analyses were run excluding those individuals who did not fully meet the inclusion criteria with regard to DQ ($n = 3$ in the experimental group; $n = 3$ in the control group; without endpoint measurement, $n = 2$) to examine whether there were any differences that would impact the study.

Results

Participant Flow

Full assessments were conducted at baseline and 1 year after the start of the intervention, which was generally about 15 months ($SD = 2.3$) after baseline (waiting time + 12 months of intervention) for both the experimental and control groups (see Fig. 1). There were no meaningful differences in results based on the 'intention to treat' or on the 'per protocol' approach. Therefore, only analyses based on 'intention to treat' are reported here.

Primary Outcomes

Table 3 shows changes in group mean scores per outcome measure before and after 1 year of intervention as well as results of the repeated measures ANOVAs. With regard to

the primary outcome, *language development*, the analyses showed no interaction effect between group and time, or between DQ, group, and time for any of the language measures, indicating no intervention effects. However, on all language measures there was a main effect of time, meaning that the language skills of children in both groups improved with time. As expected, DQ at baseline explained a significant amount of covariance in language development for all children. Analyses without those individuals who did not fully meet inclusion criteria did not influence primary outcomes.

Secondary Outcome

The change in *clinical global improvement*, as measured with the CGI-I, from baseline to endpoint was not different between the two groups [$\chi^2(2) = 0.39$, $p > 0.05$]. In the experimental group, 57% showed much improvement and 43% showed minimal or no improvement. In the control group 52% showed much improvement and 48% showed minimal or no improvement. Analyses without those individuals who did not fully meet inclusion criteria did not influence secondary outcomes.

Mediating Outcomes

Regarding *engagement*, no intervention effects were found, as represented by the non-significant group by time interactions, and these effects were not moderated by DQ at baseline. There were main effects of DQ by time, and post-hoc analyses revealed that over the 1-year period only children in the middle and higher DQ groups (≥ 50) showed an increase in compliance [$t(20) = 2.63$, $p < 0.05$; $t(11) = 3.55$, $p < 0.05$], and only children in the highest DQ group (≥ 70) showed a decrease in avoidance [$(t(11) = 3.56, p < 0.05$]. In other words, only the children in the lowest DQ group failed to show an improvement in engagement with time.

With respect to *early precursors of social communication*, we found no group by time interaction effects for any of the mediating variables, and these effects were not moderated by DQ at baseline for the ADOS outcomes. However, for Gestures (N-CDI) DQ did significantly moderate the interaction group by time. In general, there was a main effect of time on Gestures and on Social Affect (ADOS), indicating that both groups improved on these measures.

Concerning *parental skills*, no significant group by time interactions were found, neither with nor without DQ as moderator. This suggests that the mothers in the experimental group did not show an improvement in parenting skills relative to the mothers in the control group. In addition, no main effect of time was found, meaning that,

Table 3 Differences in primary and mediating outcomes from baseline to endpoint, and results of the repeated measures ANOVA's

	Experimental group (n = 36)		Control group (n = 36)		Group	Time	Group* Time	DQ	DQ* Group	DQ* Time	DQ* Group* Time
	n	Δ(EP − BL) Mean (SD)	n	Δ(EP − BL) Mean (SD)	F	F	F	F	F	F	F
Primary outcomes											
Language											
Mac Arthur—N-CDI											
Words understood	34	62.0 (75.0)	31	35.2 (66.1)	0.54	**8.19****	2.12	**33.18*****	0.02	0.72	0.29
Words said	34	75.5 (78.8)	31	56.1 (97.2)	0.00	**5.41***	2.92	**22.47*****	1.11	**4.45***	1.58
ADOS											
Level of non-echoed language on a 6 point scale	33	−1.6 (1.1)	31	−1.3 (1.2)	2.60	**27.37*****	0.26	**88.5*****	0.08	2.06	0.06
Mediating outcomes											
Engagement											
Erickson Scales											
Compliance	27	0.9 (1.5)	22	0.5 (1.5)	0.00	0.08	1.35	**7.73****	0.06	**8.98****	0.02
Non-negativity	27	0.7 (2.1)	23	0.3 (1.3)	0.95	0.99	2.10	**5.22***	0.79	0.03	1.70
Non-avoidance	27	0.7 (1.5)	22	0.5 (1.4)	0.09	0.01	0.43	**11.00****	0.09	**6.78***	0.00
Early precursors of social communication											
Mac Arthur—N-CDI											
Gestures produced	34	6.7 (10.2)	29	6.3 (9.0)	0.16	**9.76****	3.91	**17.57*****	0.02	0.05	**5.95***
ADOS											
Joint attention factor	33	−0.8 (2.3)	31	−0.9 (0.2)	0.12	0.67	0.76	**29.55*****	0.00	2.35	0.94
Social affect	33	−2.5 (4.0)	31	−2.3 (3.7)	0.01	**6.08***	0.10	**31.8*****	0.12	1.06	0.45
Parental skills											
Erickson Scales											
Supportive presence	27	0.4 (1.6)	23	0.0 (1.3)	0.04	0.34	0.73	3.21	0.02	2.64	0.00
Respect for authonomy	27	0.2 (1.2)	23	0.5 (1.3)	1.97	0.20	2.93	**4.33***	3.34	1.53	2.98
Structure and limit setting	27	1.0 (1.6)	23	0.1 (1.2)	0.54	0.02	2.52	1.00	0.07	**5.74***	0.02
Quality of instruction	27	0.5 (1.7)	23	0.0 (1.1)	0.04	1.03	0.40	2.61	0.00	**5.69***	0.44
Non-hostility	27	0.1 (0.6)	23	0.0 (0.7)	0.04	0.63	0.61	1.07	0.75	1.64	0.27

Note. DQ = Developmental Quotient; BL = baseline; EP = endpoint; ADOS = Autism Diagnostic Observation Schedule; Mac Arthur—N-CDI = Dutch version of the MacArthur Communicative Development Inventory

* $p < 0.05$; ** $p < 0.01$; ***$p < 0.001$ (printed in bold)

overall, parenting skills did not significantly improve with time in any of the groups.

Analyses without those individuals who did not fully meet inclusion criteria did influence the results for mothers' 'Structure & limit setting'. The group by time interaction on this variable became significant ($F = 4.18$, $p = 0.047$), with mothers in the experimental group improving more on this scale than mothers in the control group. No differences were found on any of the other mediating outcomes.

Discussion

In this study, we aimed to expand on previous research by evaluating the 1 year effect of a parent-training program in a comparatively large randomized controlled study, applying commonly used parent- and clinician-report based rating scales. The parent training was based on a program pilot tested by Drew et al. (2002) with promising results.

In the current study, we were unable to replicate Drew et al.'s (2002) findings. Although on most outcome variables the experimental group seemed to improve a little bit more than the control group, results were non-significant, indicating that the training program did not influence either primary (receptive and expressive language development) or secondary outcome variables (clinical global improvement). In addition, no intervention effects were found on any of the mediating outcome variables (engagement, early precursors of social communication, and parental skills), and no moderating effect of DQ was found. Although we

could not establish that the intervention had an effect, children in all DQ groups showed an improvement in language and early precursors of social communication with time, and children with a DQ of 50 or higher showed an improvement in engagement.

Several explanations may be considered for the discrepant findings between our study and that of Drew et al. (2002). First, differences in measures and informants may play a role. For example, for evaluation of verbal state, Drew et al. used the ADI-R (Rutter et al. 2003) that is based on parent judgment over a longer period, whereas we used the ADOS (Lord et al. 2001) based on judgment by a professional during one session. A second explanation may lie in variation in sample characteristics at baseline. For instance, the children in our study were about 1 year older and seemed to have a lower level of functioning than the children included in Drew et al.'s study. Third, the children in Drew et al.'s experimental group had significantly higher nonverbal IQs than the children in their control condition (88.1 vs. 66.0). Therefore, as the authors mention, it cannot be ruled out that their marginally significant finding of greater language gains in the experimental group was only due to initial difference in nonverbal IQ, and that without this difference, no effects would have been established.

Another explanation for the discrepant findings is the quality of care-as-usual. We are not in the position to judge the quality of community care in the UK (where Drew et al.'s data where obtained), but in the Netherlands care-as-usual is of a very high standard. Therefore, as both our experimental and control groups received care-as-usual, the specific added value of parent training could a priori be difficult to determine. In other words, findings in studies with a design in which an experimental training plus care-as-usual is evaluated against care-as-usual alone will always be influenced by the general standard of care-as-usual. Moreover, care-as-usual in the experimental region may have been influenced by the information provided to parents and professionals about the project prior to randomization (Oosterling et al. 2010). In the Netherlands, many centers providing special care for infants focus on 'joint attention'. As children in both groups showed improvement in child outcomes, we can conclude with appropriate caution that in the Netherlands community-based care is effective. However, to test the impact of this, one would need a design in which care-as-usual is compared to a 'no-treatment' control condition, which would not be ethical, particularly because the study period is lengthy (Lord et al. 2006).

In order to appreciate the outcomes of our negative trial, in the first place with regard to language development, we need to consider the outcomes of the mediating variables. Evidently we were not able to improve neither child nor parent mediating variables. This could represent a power

problem, but may also suggest that failure to reach the primary aims were partly caused by the failure to improve parental interaction strategies (maybe with the exception of 'Structure & limit setting'). This is crucial because the aim was to stimulate child development via their parents. So the question is what caused this. In hindsight, we think that the low frequency of the home visits might have played a role. Intervals of 6 weeks between the home visits could have been not frequent enough to achieve a result. Indeed, like Drew et al. (2002), we experienced that many parents found it difficult to consistently implement and/or maintain the recommended activities and integrate them into daily routines, often because of demands of other children, work, family life, and daily hassles. Because Drew et al. did not investigate parent–child interaction patterns, it remains difficult to compare mediating processes between studies.

Some other randomized controlled studies have included parent–child interactions in their analyses, either as a secondary or as a primary outcome measure, and did find promising results. For example, Koegel et al. (1996) studied the collateral effects of two different parent-training conditions in a randomized controlled trial ($N = 17$, age 3–9 years). One condition focused on teaching individual target behaviors (ITB) serially and used applied behavior shaping and prompting techniques. The other condition focused on pivotal response training (PRT) based on a naturalistic behavior modification approach that encourages motivation and response to multiple cues. At the end of the study, the PRT parent training approach resulted in parent–child interactions rated as happier and less stressful, and with the parents being more interested in the interaction and using a more positive communication style. In contrast, the ITB training approach did not lead to such improvements. In another randomized controlled trial applying a developmental approach ($N = 28$, age 24–72 months), Aldred et al. (2004) compared routine care alone and a dyadic nonintensive social communication training that targeted parental communication in addition to existing care. The intervention lasted 12 months. As in our study, the experimental group started with parent psycho-educational workshops, followed by individual monthly clinic sessions for 6 months with a further 6 months of 2-monthly consolidation sessions. The authors reported a significant difference in parents' observed interaction strategies with their child (in terms of greater parental synchrony) in the experimental group compared with the control group. These two studies show that it should be possible to influence parental skills, also based on non-intensive intervention (Aldred et al. 2004), but to what extent, based on what type of treatment approach (behavioral versus developmental), and dependent on which parent or child characteristics needs much more research. In addition, it is important to recognize that where there is

potential for good, there also is potential for harm. For example, the lack of improvement in their child's development, despite their intensive efforts, might adversely affect parents' self-esteem and cause stress. Therefore, it is important to carefully monitor the effects of treatments and interventions and to realize that parent training should be treated with care and may not necessarily be an essential ingredient of early intervention in all situations.

In the above, we have elaborated on possible explanations for the disappointing results. Because null-findings could occur due to Type II errors, we also need to examine the possibility of such. Type II error means not finding a between-treatment group difference when in truth there is one. First of all, a small sample size and missing values may introduce Type II error due to a power problem. Although the sample size of the current study was larger than that reported in previous parent-training studies (McConachie and Diggle 2007), an even larger sample would have improved the power of the analyses. In the current study, a sample of 68 participants was required to detect a difference between the groups. The final sample approximated this number ($N = 67$). Characteristics of outcome measures could also elevate probability of Type II error. That is, with regard to the ADOS there are three relevant issues: (a) using different modules for different children can introduce measurement error, (b) the ADOS is not specifically sensitive to very subtle differences, and (c) the psychometrics of the metric from the ADOS for the purpose of showing change in the ADOS population is unknown. In addition, the psychometrics of the CGI-I are unknown in this population for the purpose of assessing change due to treatment. The Erickson scales are quite broad and thus (subtle) differences in growth in parenting skills as a function of treatment group may be difficult to demonstrate. The issues raised with regard to ADOS, CGI-I, and the Erickson scales can all result in increased probability of Type II error. A suggestion for future intervention studies might be to perform etiologic analyses based on videotaped behavior for the exploration of more subtle differences, although the clinical relevance of differences shown by these kinds of analyses might be questioned. However, although the sample size is still relatively small and there are some limitations to the measures, we do not believe that these issues fully explain the null findings.

Although this study had several methodological strengths, such as availability of a (working) manual, inclusion of a range of widely used, blind-coded, and long-term outcomes for child development progress, inclusion of mediating and moderating factors, and application of a randomly allocated control group, it had some limitations. Despite our efforts to achieve good randomization, the study does not meet all the criteria for a perfectly designed randomized controlled trial. This could have introduced bias, but at baseline pre-treatment differences on many variables on either child or parent level that might be expected to have associations with later outcome were tested, and only a few differences between the experimental group and the control group were found. Therefore, bias based on sub-optimal randomization seems to be negligible. Furthermore, samples with more participants would allow for additional subgroup analyses. Another limitation is that we did not formally check on treatment integrity to verify if the treatment was conducted in the manner that was intended (Rogers and Vismara 2008). However, we held regular meetings during which parent-trainers discussed progress and difficulties encountered during the interventions.

To conclude, the current study sought to replicate a previous pilot randomized controlled trial of a parent-training program for preschool children with autism spectrum disorders that showed promising results. However, in our substantially larger sample these positive effects could not be replicated 1 year after the start of the intervention. Although on most outcome variables the experimental group seemed to improve a little bit more than the control group, results were non-significant, suggesting that the *Focus parent training* as performed in the context of Dutch community care did not show added value over broadly focused care-as-usual alone. In the near future, we will look into the effects of the parent training at endpoint measurement, 2 years after the start of the intervention.

References

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

Aldred, C., Green, J., & Adams, C. (2004). A new social communication intervention for children with autism: A pilot randomised controlled treatment study suggesting effectiveness. *Journal of Child Psychology and Psychiatry, 45*, 1420–1430.

Arrindell, W. A., & Ettema, J. H. M. (1975). *Handleiding bij een multidimensionele psychopathologie-indicator*. Lisse: Swets & Zeitlinger.

Bates, J. E., Freeland, C. A. B., & Lounsbury, M. L. (1979). Measurement of infant difficultness. *Child Development, 50,* 794–803.

Bayley, N. (1969). *Manual for the Bayley Scales of infant development.* New York: Psychological Corporation.

Berk, M., Ng, F., Dodd, S., Callaly, T., Campbell, S., Bernardo, M., et al. (2008). The validity of the CGI severity and improvement scales as measures of clinical effectiveness suitable for routine clinical use. *Journal of Evaluation in Clinical Practice, 14*(6), 979–983.

Bono, M. A., Daley, T., & Sigman, M. (2004). Relations among joint attention, amount of intervention, and language gain in autism. *Journal of Autism and Developmental Disorders, 34*(5), 495–505.

Bristol, M. M., Gallagher, J. J., & Holt, K. D. (1993). Maternal depressive symptoms in autism: Response to psychoeducational intervention. *Rehabilitation Psychology, 38*(1), 3–10.

Bruner, J. (1981). The social context of language acquisition. *Language & Communication, 1,* 155–178.

Charman, T. (2003). Screening and surveillance for autism spectrum disorder in research and practice. *Early Child Development and Care, 173*(4), 363–374.

Charman, T., & Baird, G. (2002). Practitioner review: Diagnosis of autism spectrum disorder in 2- and 3-year old children. *Journal of Child Psychology and Psychiatry, 43*(3), 289–305.

Charman, T., Baron-Cohen, S., Swettenham, J., Baird, G., Drew, A., & Cox, A. (2003). Predicting language outcome in infants with autism and pervasive developmental disorder. *International Journal of Language and Communication Disorders, 38*(3), 265–285.

Cicchetti, D. V., & Sparrow, S. S. (1981). Developing criteria for establishing interrater reliability of specific items: Applications to assessment of adaptive behaviour. *American Journal of Mental Deficiency, 86,* 127–137.

Dawson, G., & Osterling, J. (1997). Early intervention in autism. In M. Guralnick (Ed.), *The effectiveness of early intervention.* Balimore, MD: Brookes.

Dawson, G., Toth, K., Abbott, R., Osterling, J., Munson, J., Estes, A., et al. (2004). Early social attention impairments in autism: social orienting, joint attention, attention to distress. *Developmental Psychology, 40*(2), 271–283.

de Brock, A. J. L. L., Vermulst, A. A., Gerris, J. R. M., & Abidin, R. R. (1992). *NOSI, handleiding experimentele versie.* Amsterdam: Pearson.

Drew, A., Baird, G., Baron-Cohen, S., Cox, A., Slonims, V., Wheelwright, S., et al. (2002). A pilot-randomised control trial of a parent training intervention for pre-school children with autism. Preliminary findings and methodological challenges. *European Child and Adolescent Psychiatry, 11,* 266–272.

Erickson, M. F., Stroufe, L. A., & Egeland, B. (1985). The relationship between quality of attachment and behavior problems in preschool in a high-risk sample. *Monographs of the Society for Research in Child Development, 50,* 147–166.

Fenson, L., Dale, P. S., Reznick, J. S., Thal, D., Bates, E., Hartung, J. P., et al. (1993). *The MacArthur Communicative Developmental Inventories: User's guide and technical manual.* Baltimore: Paul H. Brookes Publishing Co.

Gotham, K., Risi, S., Pickles, A., & Lord, C. (2007). The autism diagnostic observation schedule: Revised algorithms for improved diagnostic validity. *Journal of Autism and Developmental Disorders, 37,* 613–627.

Guy, W. (1976). *ECDEU assessment manual for psychopharmacology.* Rockville, MD: U.S. Department of Health, Education, and Welfare.

Jocelyn, L. J., Casiro, O. G., Beattie, D., Bow, J., & Kneisz, J. (1998). Treatment of children with autism: A randomized controlled trial to evaluate a caregiver-based intervention program in community daycare centers. *Journal of Developmental and Behavioral Pediatrics, 19*(5), 326–334.

Koegel, R. L., Bimbela, A., & Schreibman, L. (1996). Collateral effects of parent training on family interactions. *Journal of Autism and Developmental Disorders, 26*(3), 347–359.

Lord, C., Rutter, M., DiLavore, P., & Risi, S. (2001). *Autism Diagnostic Observation Schedule (ADOS) manual.* Los Angeles, CA: Western Psychological Services.

Lord, C., Wagner, A., Rogers, S., Szatmari, P., Aman, M., Charman, T., et al. (2006). Challenges in evaluating psychosocial interventions for autistic spectrum disorders. *Journal of Autism and Developmental Disorders, 35*(6), 695–708.

McConachie, H., & Diggle, T. (2007). Parent implemented early intervention for young children with autism spectrum disorder: A systematic review. *Journal of Evaluation in Clinical Practice, 13,* 120–129.

Mullen, E. (1995). *Mullen scales of early learning* (AGS ed.). Circle Pines, MN: American Guidance Service.

Mundy, P., Sigman, M., & Kasari, C. (1990). A longitudinal study of joint attention and language development in autistic children. *Journal of Autism and Developmental Disorders, 20,* 115–128.

National Research Counsil. (2001). *Educating children with autism.* Washington, DC: National Academy Press.

Oosterling, I. J., Swinkels, S., Van der Gaag, R. J., Visser, J. C., Dietz, C., & Buitelaar, J. (2009). Comparative analysis of three screening instruments for autism spectrum disorder in toddlers at high risk. *Journal of Autism and Developmental Disorders, 39,* 897–909.

Oosterling, I. J., Wensing, M., Swinkels, S. H., van der Gaag, R. J., Visser, J. C., Woudenberg, T., et al. (2010). Advancing early detection of autism spectrum disorder by applying a two-stage screening approach. *Journal of Child Psychology and Psychiatry, 51*(3), 250–258.

Ospina, M. B., Krebs Seida, J., Clark, B., Karkhaneh, M., Hartling, L., Tjosvold, L., et al. (2008). Behavioural and developmental interventions for autism spectrum disorder: a clinical systematic review. *PLoS ONE, 3*(11), e3755.

Rogers, S. J., & Vismara, L. A. (2008). Evidence-based comprehensive treatments for early autism. *Journal of Clinical Child Psychology, 3*(1), 8–38.

Rutter, M., Le Couteur, A., & Lord, C. (2003). *ADI-R: Autism diagnostic interview—revised.* Los Angeles, CA: Western Psychological Services.

Schopler, E., Reichler, R. J., Bashford, A., Lansing, M. D., & Marcus, L. M. (1990). *The psychoeducational profile revised (PEP-R).* Austin, TX: Pro-Ed.

Siller, M., & Sigman, M. (2002). The behaviors of parents of children with autism predict the subsequent development of their children's communication. *Journal of Autism and Developmental Disorders, 32*(2), 77–89.

Swinkels, S. H., Dietz, C., Daalen, E., van Kerkhof, I. H., van Engeland, H., & Buitelaar, J. K. (2006). Screening for autistic spectrum in children aged 14 to 15 months I: The development of the Early Screening of Autistic Traits Questionnaire (ESAT). *Journal of Autism and Developmental Disorders, 36*(6), 723–732.

Toth, K., Munson, J., Meltzoff, A. N., & Dawson, G. (2006). Early predictors of communication development in young children with autism spectrum disorder: Joint attention, imitation and toy play. *Journal of Autism and Developmental Disorders, 36,* 993–1005.

Whalen, C., Schreibman, L., & Ingersoll, B. (2006). The collateral effects of joint attention training on social initiation, positive affect, imitation, and spontaneous speech for young children with autism. *Journal of Autism and Developmental Disorders, 36,* 655–664.

Zink, I., & Lejaegere, M. (2002). *N-CDI: Lijsten voor communicatieve ontwikkeling.* Leuven: Acco.

Zink, I., & Lembrecht, D. (2000). *Nederlandstalige non-speech test.* Leusden: Acco.