Journal of Autism and Developmental Disorders (2020) 50:844–863
https://doi.org/10.1007/s10803-019-04316-2

ORIGINAL PAPER

Preschool Peer Social Intervention (PPSI) to Enhance Social Play, Interaction, and Conversation: Study Outcomes

Nirit Bauminger-Zviely[1] · Dganit Eytan[1] · Sagit Hoshmand[1] · Ofira Rajwan Ben–Shlomo[1]

Published online: 27 November 2019

Abstract

This RCT study examined efficacy of a preschool peer social intervention (PPSI) in facilitating social engagement of preschoolers with high-functioning ASD (HFASD; $N=65$). HFASD participants were randomly assigned by preschool to a 6-month intervention (play, interaction, or conversation) or a waitlisted-treatment-as-usual control group. Trained on-site therapists led the PPSI in preschools, in small ($n=3$–4) mixed (HFASD/typical) groups. Results showed that all intervention groups improved over time, each mainly in its own targeted peer-engagement domain, but the control group even deteriorated on some measures. Intervention groups also showed generalization to untrained domains (adaptive skills) and settings (play complexity during preschool activities). It is advised that individualized needs-based holistic peer intervention, comprising all three domains, should be part of early ASD intervention.

Keywords High-functioning children with autism spectrum disorders (HFASD) · Social intervention · Peer interaction · Social play · Peer talk and conversation

Peer relations are cardinal for children's development of ample cognitive, linguistic, and social skills (e.g., Hay et al. 2009). Longitudinal evidence shows that individual variations in behavior and in responding to peers' behavior at early ages predict later social competence (e.g., Hay et al. 2009). Also, lower levels of peer engagement in early childhood have been linked with indicators of poorer social adaptation in the general population (Vaughn et al. 2016). Yet, in the three key domains associated with efficient peer interaction—namely, social interactive skills, play, and conversation (Coplan and Arbeau 2009)—young children with autism spectrum disorder (ASD) do not appear to show the same patterns as their typically developing (TYP) counterparts (e.g., Jordan 2003; Stefanatos and Baron 2011). The early social engagement challenges experienced by young children with ASD may limit their early peer relationship experiences and reduce later peer engagement across development (Manning and Wainwright 2010; Schuler and Wolfberg 2000).

Although peer relations comprise a core social-communicative deficit in ASD (e.g., APA 2013) relevant randomized controlled trials (RCTs) of peer programs for young children with ASD are scarce. The current study aimed to close this research gap by evaluating the efficacy of the *preschool peer social intervention* (PPSI) in facilitating peer engagement among preschoolers with ASD.

Social Interaction in ASD and TYP

Social interaction is defined as a reciprocal process in which children effectively initiate and respond to social stimuli presented by their peers in diverse social settings and situations (Shores 1987). Abilities like joining a group, responding to others, coping with conflict situations, and demonstrating prosocial behaviors such as offering help and comfort are examples of the joint interaction skills that are important for developing adaptive peer engagement (Coplan and Arbeau 2009). In the TYP population, peer interactions gradually become mutual and coordinated during the preschool years (Hay et al. 2009).

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10803-019-04316-2) contains supplementary material, which is available to authorized users.

✉ Nirit Bauminger-Zviely
nirit.bauminger@biu.ac.il

[1] School of Education, Bar-Ilan University, 529002 Ramat Gan, Israel

Research has shown that children with ASD tend to have infrequent and low-quality peer interactions (i.e., behaviors that indicate social intention but with minimal social enactment; see review in Bauminger-Zviely 2013). Among some of the low-quality behaviors that children with ASD may demonstrate are close proximity, vague gaze, imitation of others' social acts, functional behaviors (e.g., giving and requesting information), passive observation, and solitary behaviors (see review in Bauminger-Zviely 2013). The production of complex interactive prosocial behaviors such as sharing, providing help, expressing positive affect, making eye contact, cooperating, and comforting is often challenging for many children with ASD. In older children with high-functioning ASD (HFASD; IQ > 70), responding to peer interaction is more intact than initiating interaction. However, such responses' quality may be poor (active but odd), which can preclude the development and maintenance of fruitful or extended social interactions with peers. To be noted, most available knowledge to date is from observational data of older children and low-functioning children with ASD. Nevertheless, a recent study from Bauminger's research laboratory also reported lower peer interaction capabilities in preschoolers with HFASD versus preschoolers with TYP, substantiating former data. Greater difficulties emerged in these preschoolers' prosocial behaviors, social conversation, and social and social pretend forms of play, whereas functional communication was more intact (Bauminger-Zviely and Shefer 2019).

Social Play in TYP and ASD

Social forms of play involve active and enjoyable engagement on the part of at least two players. Such play is characterized as spontaneous and voluntary. It is also free from means-end directiveness (i.e., playing is an end in and of itself); is nonliteral (i.e., imaginative rather than strictly realistic); and should be flexible and dynamic (e.g., Bauminger-Zviely 2013). During toddlerhood and the preschool years, children with TYP gradually develop their capabilities for both social play and social pretend play to form complex, reciprocated peer play, based on peers' shared concrete and abstract play acts (Coplan and Arbeau 2009; Howes 2011).

In social play, toddlers engage in parallel play activities in close proximity to peers. Young children (1.5 to 2.5 years) begin simple social play that involves direct social behavior with peers. Next (2.5 to 3 years), they move on to interactive-complementary play skills, including role reversals in social games such as run-and-chase or block building. During the preschool years (3–5 years), such play becomes reciprocal and involves joint planning so that the pair's actions are integrated (Howes 1980; Howes et al. 1992).

In social pretend play, children learn to share meanings of imaginary social acts and to negotiate and resolve conflicts. In the second year, children move from engaging in solitary pretend acts to those directed toward a partner through simple unnamed scripts, with or without complementary roles (e.g., bus driver and passenger). During the preschool years (3–5), children learn to engage in complex social pretend play during which they meta-communicate over roles, scripts, and themes ("Let's play doctor; I will be the doctor and you will be the patient").

Without explicit guidance, children with autism are frequently at risk for being deprived of consistent social play and social pretend play experiences. Compared to age-mates with TYP, their play often manifests itself in less diverse and complex forms and sometimes idiosyncratically (e.g., Campbell et al. 2017; Hobson et al. 2013; Jordan 2003; Wolfberg 2016). The symbolic dimension of pretend play is often a major challenge in ASD. Many of these children may exhibit mere repetition of a peer's pretending actions, fewer novel play acts, and less elaboration on pretend play, compared to age-mates with TYP (e.g., Murdock and Hobbs 2011). Particularly challenging for many children with ASD are free-play situations involving peers. Difficulties in co-regulating one's own play behaviors with others' has been highlighted as a cardinal issue in the development of social play in many young children with ASD (Jordan 2003).

Social Conversation in TYP and ASD

Peer talk is crucial for pragmatic and social development. It offers children a wide range of opportunities for mutual learning of social-interactive as well as linguistic skills (Blum-Kulka and Snow 2004). The situational embeddedness of preschoolers' talk varies greatly, differing in its degree of connection to the activity at hand ("activity talk") and in its conversational genre such as interpersonal sharing of personal experience and feeling or argumentative discourse while discussing a topic from different viewpoints (e.g., Blum-Kulka et al. 2004). Children with TYP as young as preschool ages are able to demonstrate fairly complex forms of peer talk that are directed toward the listener.

The pragmatic deficit, reflected in deficient conversational capabilities, is considered the hallmark of the language deficit in ASD (see review in Stefanatos and Baron 2011). Although spontaneous peer talk in ASD has rarely been examined during preschool, two recent studies presented such data. In Bauminger-Zviely et al. (2014), peer talk in preschoolers with HFASD versus preschoolers with TYP was examined during a 10-min free-play/snack time. The main results demonstrated a more intact pragmatic profile in TYP than in HFASD, with the HFASD group showing the greatest difficulty on several Pragmatic Rating Scale-Young

(Bauminger-Zviely et al. 2014) subscales. Namely, they demonstrated little reciprocal conversation, unresponsiveness to conversational partners, out-of-context utterances (on the "pragmatic" subscale); unusual intonation and stereotypic speech ("prosodic" subscale); and unusual eye contact, inappropriate facial expression, and inappropriate gesture ("paralinguistic" subscale). Bauminger-Zviely et al. (2017) compared HFASD and TYP groups for children's speech acts (i.e., communicative intentions conveyed by utterances in specific contexts) during peer talk in a free-play situation. A higher number and diversity of speech acts emerged in TYP than in HFASD, mainly in actively/assertively reporting facts, evaluating situations/conditions, stating rules, conveying attitudes, and using organizational devices. The latter meta-linguistic acts are used to regulate contact and conversational flow.

These relatively rare data on young children with ASD corroborate prior data on poor social conversational skills among older children with ASD. What seem to be especially difficult are some of these children's abilities to initiate, develop, or expand an interaction by taking turns within an ongoing conversation or by switching between topics to accommodate the conversational partner's perspective. Older children's talk has been reported as often focused on their own areas of interest, with frequent difficulty conversing about the partner's interests, and at times some children may leave conversations without coherently ending them (Müller et al. 2016; Paul et al. 2009, 2014; Rubin and Lennon 2004). These pragmatic difficulties in conversational ability seem likely to exert an impact on preschoolers' ability to take part in productive peer interactions.

Prior Interventions to Facilitate Peer Interaction, Play, and Conversation in ASD

Although specific, focused, child–child interventions are clearly needed to enhance the peer interactions, social forms of play, and social conversation among young children with ASD, interventions promoting these key domains at preschool ages have not yet received sufficient empirical attention. Reichow and Volkmar (2010) found that only 14% of social skills interventions during preschool involved peers as delivery agents. Likewise, in reviewing 14 articles that examined peer-mediated interventions' efficacy in ASD, Watkins et al. (2015) found that only 3 articles (21%) referred to preschoolers ages 4 to 5 years. Several ASD studies have revealed improvement in complex forms of play and in conversational skills but mainly through child–adult interactions (e.g., Chang et al. 2016; Conallen and Reed 2017; Koegel et al. 2014; Wilson et al. 2017). Indeed, most studies to date that attempted to promote the conversational, interactive, and/or play skills of children with ASD had different foci, such as older school-age children/adolescents rather than preschoolers. Also, adult–child engagement was more often at focus than naturalistic child–child engagement. Furthermore, some studies were methodologically limited by small sample sizes, the absence of randomization and control groups, and/or lack of simultaneous separate delivery of interventions in all three key domains within a comprehensive holistic model (e.g., Luckett et al. 2007; Müller et al. 2016; Reichow and Volkmar 2010; Rogers 2000; Thiemann and Goldstein 2004).

Despite the limited scope of prior peer interventions during preschool, the studies that did implement such peer-engagement procedures revealed positive results. For example, peer-mediated interventions utilizing peers with TYP as mediators, to enhance social interactive skills among preschoolers with ASD, resulted in increases in their initiations and responses to peers (e.g., Kalyva and Avramidis 2005; Kamps et al. 2015; Katz and Girolametto 2015). Some play-targeted peer interventions were structured and included pre-training of peers with TYP or siblings (Oppenheim-Leaf et al. 2012; Neff et al. 2017) or utilized peers with TYP as role models to enhance the social interaction and play behaviors of preschoolers with ASD (e.g., Ganz and Flores 2008). Others used scripts and role-play to increase peer interaction and play behaviors during interactive play groups (e.g., Murdock and Hobbs 2011) or during structured outdoor activity (e.g., Morrier and Ziegler 2018). Some studies targeted play behaviors in small mixed groups by offering naturalistic free-play areas containing a range of developmentally and age-appropriate play materials with high social and imaginative potential, such as Wolfberg's (2016) integrated play group model.

In all, the findings of these peer-oriented interventions demonstrated increases in play complexity, play and interaction adequacy, and even generalization to situations with unfamiliar peers (e.g., Watkins et al. 2019). In addition, an advantage was noted for peer mediation versus adult mediation (Strauss et al. 2014). Integrating peers with TYP into social intervention programs seems essential when the goal is to promote social interactive skills. Overall, it appears that the ASD field could benefit from more evidence-based RCTs investigating manualized peer interventions in naturalistic settings—targeting spontaneous peer interaction, social play and social pretend play, and social conversation—to contribute to improvements in these preschoolers' quality of life by mitigating social-communicative symptoms associated with disability.

The Current Study

This study's major aim was to conduct an RCT to examine the efficacy of a novel, manualized, preschool-based, child–child intervention (the PPSI) for increasing the ability

of preschoolers with HFASD to socially engage (play, interact, and converse) with peers. The PPSI's design espoused a holistic conceptual view of peer engagement encompassing three key domains: play, interaction, and conversation. Thus, each of three preschooler groups received a different intervention: the "social interaction" (INTERACT), the "social conversation" (CONVERSE), and the "social/social-pretend play" (PLAY) intervention groups. A fourth group, the waitlisted (WAIT) control group, received treatment as usual while awaiting intervention.

The PPSI's psychoeducational program in each of the three intervention domains comprised a transactional approach incorporating developmental features, principles and methods taken from cognitive-behavioral therapy, and a naturalistic ecological orientation. Developmentally, this novel PPSI curriculum included age-appropriate play, interaction, and conversation skills while accounting for individual differences in social communication abilities (Krasny et al. 2003). Due to the young age of our participants, a full cognitive-behavioral therapy model could not be performed; however, we used age-appropriate adaptations such as cognitive techniques like problem solving and concept clarification for the learning of social constructs, combined with behavioral techniques like role-play for practicing peer engagement behavior within the peer group (Dobson and Dobson 2009). Ecologically, the intervention was implemented within the participants' natural social environment (i.e., preschool) and included children's social agents, namely, their peers and teachers (Bronfenbrenner 1992).

More specifically, this study's major questions were to examine if each treatment program would be effective in facilitating preschoolers' peer engagement (pre–post differences), if peer engagement would differ after treatment between the four groups, and if trained skills would be generalized to other settings (i.e., everyday preschool activities) and other contents (e.g., untrained domains; adaptive skills). Thus, we aimed to examine the PPSI's efficacy by evaluating changes in the targeted preschoolers' social engagement before versus after treatment, in each of three groups receiving a different intervention (INTERACT, CONVERSE, PLAY). We assumed that comparing these three key treatment modalities with one another would help delineate each modality's contribution to improvements in social engagement among participants with HFASD. Our prediction was that the targeted children with HFASD in each of the three intervention groups would improve over time in their trained social engagement domain. Namely, we expected increased conversation quality in the CONVERSE group, increased play complexity in the PLAY group, and some improvement on both the conversation and play scales for the INTERACT group.

With regard to generalization, we predicted that the preschoolers would improve in spontaneously applying trained skills while engaging with other peers in untrained non-PPSI contexts (outside the trained social group) during everyday preschool activities. Based on some limited prior research with older children, we also speculated that the participating preschoolers would show some generalization to untrained areas of social functioning (adaptive skills and socialization). In all these areas, we expected the three treatment groups to improve to a greater degree than children in the WAIT group.

Method

Participants and Procedure

Sixty-five children with ASD (four girls) ages 4.0–5.2 years participated in the study. Inclusion criteria were: (1) ASD diagnosis according to the Autism Diagnostic Observation Schedule-Generic (ADOS-G; Lord et al. 2000), which verified previous clinical diagnosis by licensed psychologists unassociated with the current study; (2) no comorbid diagnosis (e.g., ADHD, language disorder, sensory or motor disabilities); and (3) IQ > 70 according to the Mullen Scales of Early Learning (Mullen 1997) to denote sufficiently high cognitive and language abilities to participate in the intervention. To be noted, we defined our participants with ASD as high-functioning based on their IQ level beyond intellectual disability. This more conservative definition of functioning does not ignore recent discussion in the field striving to reach a more exhaustive definition that goes beyond cognitive abilities and includes other areas of functioning such as adaptive behaviors (e.g., Alvares et al. 2019).

Recruitment Process

Children with ASD were recruited from 23 different special education preschools located in middle-class, large, urban areas throughout Israel. Over 2 years of recruitment, 74 candidates with ASD were identified, but nine were excluded from participation (see Fig. 1). Altogether, the study included 65 children with ASD, 43 recruited in the first year and 22 in the second year, attending 23 special education preschools. We then recruited children with TYP to regularly meet with their peers in small groups. The latter were recruited from regular preschools nearby the targeted special preschools, based on regular preschool educators' reports that these children had no identified disability of any kind, similar chronological age as the nearby target children with ASD, and adequate social and behavioral capabilities, excluding children with behavioral difficulties. Data were collected only on the targeted children with ASD due to Ministry of Education policy. All permissions and informed

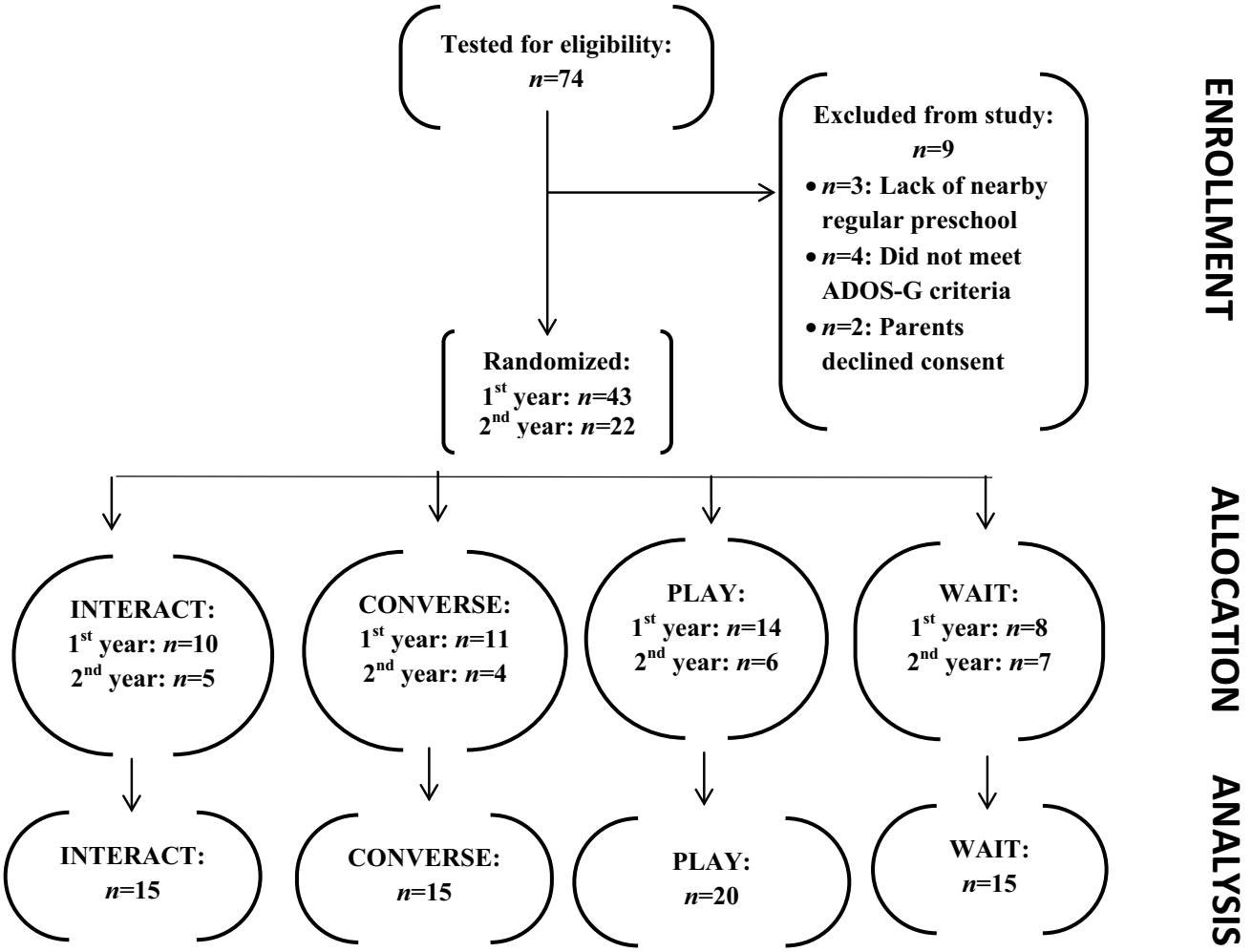

Fig. 1 CONSORT flow chart illustrating sample selection of preschoolers with ASD

consents were obtained from governmental officials, preschool personnel, parents, and children.

Randomization

Using a randomized block design, the 23 participating special education preschools were randomly assigned to the four study groups (CONVERSE, INTERACT, PLAY, or WAIT) along the two recruitment years, including a division between three geographical regions in proportion to the number of children with ASD and special education preschools recruited per region. Thus, all four study groups were allocated to each of three regions, comprising nine preschools in northern Israel ($n=24$ children: PLAY-6; CONVERSE-8; INTERACT-6; WAIT-4); eight preschools in southern Israel ($n=23$ children: PLAY-5; CONVERSE-5; INTERACT-7; WAIT-6); and six preschools in central Israel ($n=18$ children: PLAY-9; CONVERSE-2; INTERACT-2; WAIT-5). Each preschool received only one type of intervention, provided by only one trained therapist, who led 1–3 study groups over the 2-year period, totaling 28 small groups. Altogether, all children with ASD meeting the inclusion criteria ($n=65$; four girls) were assigned to one of three intervention groups—INTERACT ($n=15$), PLAY ($n=20$, two girls), or CONVERSE ($n=15$, 2 girls)—or to the WAIT treatment-as-usual control group ($n=15$), based on their randomly allocated preschool. As seen in Table 1, the four study groups of children with ASD did not differ significantly on chronological age, IQ scores, ADOS-G scores, or mothers' education.

The PPSI Intervention

Setting/Structure

Trained on-site therapists in the special education preschool led the manualized intervention over a 6-month period in three 45-min sessions per week held in a quiet separate

Table 1 Sample characteristics of preschoolers with HFASD by group

	INTERACT ($n=15$)	CONVERSE ($n=15$)	PLAY ($n=20$)	WAIT ($n=15$)	Group Dif. F (3, 61)
CA (months)					
M(SD)	52.26 (7.20)	58.53 (6.88)	54.45 (8.91)	56.60 (7.27)	1.86
Range	48.27–56.26	54.54–62.53	50.99–57.90	52.60–60.59	
IQ[a]					
M(SD)	98.80 (17.03)	96.13 (17.41)	96.50 (14.66)	103.40 (13.18)	0.72
Range	90.74–106.85	88.08–104.18	89.52–103.47	95.34–111.45	
Verbal IQ[a]	97.40 (16.80)	94.66 (17.95)	95.60 (18.26)	100.40 (14.28)	0.34
	88.61–106.18	85.88–103.45	87.99–103.20	91.61–109.18	
Nonverbal IQ[a]					
M(SD)	100.33 (21.19)	95.73 (19.98)	96.65 (16.10)	105.73 (15.48)	0.98
Range	90.94–109.72	86.34–105.12	88.52–104.77	96.34–115.12	
ADOS-G[b]					
M(SD)	10.40 (2.19)	9.40 (1.72)	10.68 (2.26)	11.26 (3.21)	1.59
Range	9.15–11.64	8.15–10.64	9.58–11.78	10.02–12.50	
Mother's education[c]					
M(SD)	4.46 (0.83)	4.60 (1.18)	5.05 (0.99)	4.33 (0.97)	1.73
Range	3.94–4.98	4.08–5.12	4.60–5.49	3.81–4.85	

[a]Mullen Scales of Early Learning

[b]Autism Diagnostic Observation Schedule-Generic social interaction + communication

[c]Education calculated on a 6-point scale from elementary school (1) to master's degree and up (6)

room. One session per week comprised an experiential learning activity for that week's social construct, such as short video clips or role-playing puppets demonstrating the learned target behavior. The learning session was delivered to the children with ASD in their special education setting. The other two sessions per week enabled practice of that week's learned social construct in small mixed peer groups ($n=3$–4) consisting of one or two children with HFASD and two peers with TYP. Each week's target behavior was practiced via peer group play with toys, joint activity with crafts, maintaining conversation while eating snacks, and so on. The same children with TYP participated in all practice sessions across the intervention, in line with former studies in the field showing that children with ASD show better social functioning with familiar than unfamiliar peers (Bauminger-Zviely 2013). Familiarity enables the development of rapport during group work and thus may lead to better treatment results. The 2 weekly practice sessions were held consistently in a single setting: Of the 28 mixed small groups, 17 met in the special education preschool and 11 met in the regular preschool.

Intervention Content

The PPSI intervention curricula were chosen based on typical play, interaction, and conversation development, and in line with the social and communication deficits found to characterize young children with HFASD in these domains, as described above. Specifically, the INTERACT curriculum was adapted from Bauminger's (2002, 2007a, b) cognitive-behavioral ecological model for school-age children. It focused on social interaction skills (e.g., joining a group; implementing conflict resolution; prosocial skills like encouraging, helping, comforting, expressing positive affect). The PLAY curriculum followed Howes's (1980) Howes and Matheson 1992) stages of *social play* (parallel, parallel aware, simple, complementary, and reciprocal) and of *social pretend play* (isolated pretend acts, cooperative, and complex involving meta-communication about play). The CONVERSE curriculum focused on social conversation stages (e.g., initiating, developing, ending) and age-appropriate conversational genres, including interpersonal conversation, argumentative discourse, and activity talk (Blum-Kulka et al. 2004). At the start of all three interventions, preschoolers received instruction regarding basic language concepts essential for all treatments (e.g., definition of friend/group; same/different; sharing; listening). Appendix A (Supplementary Material) presents some sample session components for each of the three intervention groups.

Intervention Techniques

We used various techniques throughout the learning and practice sessions including: concept clarification; problem solving; role-play; doll play; various visual or audiovisual

stimuli like illustrations, photos, and short video clips; fun play activities like various games and creative crafts.

The Role of Peers with TYP

Children with TYP were asked to participate in all group activities as social partners. Their participation provided optimal modeling for adapting social behaviors to varying social situations. They were asked to participate in all group activities and engage with the other children in the group. They were not provided with specific instructions other than their cooperation around the sets of activities organized by the therapists, according to the various social goals to converse, interact, or play, with an emphasis on fun group activities.

Intervention Fidelity

Therapists underwent a full-day training seminar at the university in three separate groups focusing on their assigned INTERACT, CONVERSE, or PLAY intervention. All intervention groups were supervised in preschools every 2 weeks by the research team to ensure uniformity and consistency in program delivery according to the manualized protocol. The research team monitored the intervention implementation by observing the sessions, noting any deviations from the assigned protocol, and providing feedback and guidance to therapists. In addition, each therapist also completed a 2-page fidelity questionnaire once weekly (e.g., Skolits and Richards 2010; Stein et al. 2008), reporting any changes they made in implementing five components of the intervention (equipment, activities' arrangement and duration, instructions given to children, and other changes) and why. One page referred to that week's social construct learning activity, and the other page referred to the ensuing two mixed-group practice sessions. The research team reviewed and discussed these weekly fidelity questionnaires with each therapist in their biweekly fidelity meetings, to monitor therapists' adherence to the assigned protocol. Fidelity percentages were calculated for each week and then averaged for the entire program, yielding the following ranges: INTERACT—91 to 98% ($M = 93.3$), CONVERSE—86 to 100% ($M = 97.0$), and PLAY—82.5 to 100% ($M = 96.01$). Finally, each therapist completed a self-reported fidelity summary questionnaire at the end of the intervention program. The summary report included 9 statements (e.g., "I conducted the program in the most precise way according to the protocol") rated on a 5-point Likert scale (1 = *do not agree at all* to 5 = *strongly agrees*), with a Cronbach α of 0.72. Mean fidelity summary scores were high for all groups: INTERACT—4.04 ($SD = 0.60$), CONVERSE—4.30 ($SD = 0.45$), PLAY—4.74 ($SD = 0.43$).

Assessment Measures

All assessments were collected both at Time1 (T1—baseline) and Time2 (T2—immediately after 6-month treatment) for all four groups including the control group. Triangulation of data sources (direct observations, assigned therapist's ratings, and uninvolved head educator's ratings) measured the targeted PPSI intervention's impact.

Direct Observations to Examine T1–T2 Change in Targeted Interactive Play and Conversation Behaviors as Manifested in Social Engagement During Free Time Within Assigned Groups

Pre–post improvement in trained areas was evaluated by direct observation of participants' spontaneous social engagement during free-time situations. These free-play and snack-time situations with assigned mixed-group members were videotaped and coded to assess participants' spontaneous display of the targeted interactive conversational and play behaviors with the same group of peers but outside of directly mediated PPSI activities.

Observed Hierarchical Stage of Play Development Based on Howes's (1980) and Howes and Matheson's (1992) Peer Play Scale assessing the complexity of peer play, participants' hierarchical-developmental stages of both social play and social pretend play were coded. Coding was based on a videotaped 15-min free-play situation within their assigned small group. Groups received the same toy set at both time points (T1, T2), including abstract toys (e.g., rings, sticks, blocks) and concrete toys such as dolls, animals, birthday party items (e.g., candles, cake), bath items (e.g., mini hairbrush, shampoo, bathtub), kitchen items (e.g., mini cooking pot, cutlery, vegetables, fruits), doctor toys (e.g., plastic stethoscope, thermometer, bandage), and transportation toys (e.g., cars, traffic signs). Coders marked the highest social play stage and the highest social pretend play stage achieved by the participant during the 15-min observation. Each scale was rated 1–5, representing a developmental sequence from non-interactive parallel play with no social interaction, to contingent social interaction, to reciprocal and complementary interaction (see Appendix B in Supplementary Material). Coders were masked to study aims and to participants' group assignments. Interrater agreement was calculated between the masked observer and the third author on 25% of observations, randomly divided between study groups. This yielded 90.00% agreement on the social play scale and 94.44% on the social pretend play scale. This scale has previously been used successfully in the study of ASD (e.g., Sigman and Ruskin 1999).

Observed Conversation Pragmatics and Quality Participants' conversation capabilities within their assigned small groups were evaluated during two videotaped social situations: the 15-min free-play session described above and a 10-min snack time when the group received snacks and a set of toys (plastic food items, transportation toys, animals, and plastic fences). Due to the lack of a significant situation effect, we combined both situations (totaling 25 min.) in our pragmatics and conversational quality analyses. Participants' conversational acts were transcribed verbatim by a speech therapist and coded using two scales, as follows. For each of these two scales, two observers (the fourth author and a trained assistant, both of whom are senior speech therapists and experts in ASD) coded a randomly selected sample of 25% of the participants along the various groups and times. They obtained interrater agreement of 90.90% for the pragmatic scale and 84.58% for the quality scale. The second coder then worked independently on the remaining data.

Observed Pragmatic Performance During Conversational Acts The Pragmatic Rating Scale-Young (PRS-Y; Baumginger-Zviely et al. 2014) was utilized to code participants' videotaped free-play and snack-time social situations to assess 27 pragmatic behaviors that reflect abnormalities reported to be characteristic of autism, based on theoretical and clinical reports of major pragmatic behaviors in the literature (Landa et al. 1992). At the end of each situation (free-play/snack), each PRS-Y behavior was rated on a 3-point scale: 0 = *almost never occurs*; 1 = *occurs sometimes but does not seriously interfere with conversational flow*; and 2 = *occurs almost always, seriously interfering with conversational flow*. Higher scores (the mean for the entire 25-min interaction) indicated a more severe pragmatic deficit. The PRS-Y items comprised three subscales: (1) *pragmatic behaviors*, 15 items on topic management and reciprocity (e.g., "overly talkative," "unresponsive," "vague;" $\alpha = 0.79$); (2) *speech and prosodic behaviors*, 8 items on the speaker's production (e.g., "indistinct speech," "unusual intonation," "unusual rhythm;" $\alpha = 0.61$); and (3) *paralinguistic behaviors*, 4 items on physical behaviors that accompany speech (e.g., gestures, facial expressions, physical distance, gaze; $\alpha = 0.79$).

Observed Conversation Quality Participants' videotaped free-play and snack-time social situations were assessed for the five major conversational capabilities previously identified as significant social conversation components (e.g., Adams 2002; Paul et al. 2009): (a) *attention*—behaviors to obtain listener's attention (e.g., calling his/her name); (b) *initiation*—behaviors to initiate new relevant or related content; (c) *responding*—behaviors to appropriately join an existing conversation (e.g., asking questions); (d) *preservation*—behaviors to preserve/maintain an existing conversational topic; and (e) *expansion* of current conversational topic (e.g., adding relevant information). Each participant's conversational turn (utterance) was coded on a 5-point scale from 1 (*one or fewer conversational components*) to 5 (*inclusion of all 5 conversational components*). A mean score was calculated across both social situations to signify the quality of the participant's observed conversation, with higher scores indicating better quality ($\alpha = 0.93$).

Informant Ratings to Examine T1–T2 Change in Targeted Interactive-Play and Conversation Skills as Manifested in Everyday Social Engagement in the Preschool

Children's pre–post improvement in trained skills was also evaluated by two informants' questionnaires. The assigned intervention therapist and the preschool's uninvolved head educator each assessed participants' spontaneous display of the targeted interactive conversational and play behaviors as manifested in everyday non-mediated indoor and outdoor peer social situations in the preschool. In addition, to assess generalization of trained skills to untrained domains and settings, we traced pre–post changes in the targeted preschoolers' spontaneous display of skills from the other two treatment domains that had not been trained as manifested in everyday preschool activities (e.g., the CONVERSE group's generalization from trained language-conversation skills to untrained play skills in non-PPSI preschool social situations). Although both informants rated participants' social engagement at pretest and posttest, only the uninvolved head educators were regarded as full informants on generalization because they were masked to intervention protocols and to the evaluated children's intervention group assignment. The therapists were an important additional data informant to complement the direct observations in order to assess spontaneous peer engagement in the preschool beyond the mediated PPSI activities.

Therapist-Rated Play Skills The Social Play Questionnaire (SPQ) was developed for the purpose of the current study to tap therapists' evaluation of the social play and social pretend play stages that their assigned participant(s) exhibited during spontaneous non-mediated social situations in the preschool (outside the study group). The 43-item SPQ, based on Howes's (1980) and Howes and Matheson's (1992) Peer Play Scale, therefore yielded 10 developmental stages, 5 each for the social play and social pretend play scales (see Appendix B for developmental sequence, sample items, and reliabilities). Items were rated on a 5-point Likert scale from 1 (*does not exist at all*) to 5 (*exists consistently in participant's play repertoire*). Both for social play and social pretend play, pre-analyses yielded only three developmental stages due to mid-to-high correlations (0.44–0.93) between

Stages 2 and 3 and Stages 4 and 5. Thus, all further analyses utilized only the three developmental levels (Stages 1, 2–3, and 4–5) for both the social play and social pretend play scales (see grey columns in Online Appendix B).

Therapist-Rated Conversation Skills The Social Conversation Questionnaire (SCQ) was developed for the purpose of the current study to tap therapists' evaluation of the peer social conversation skills that their assigned participant(s) exhibited spontaneously during non-mediated social situations in the preschool. Development of the 23-item SCQ-T was grounded in the main conversational capabilities that were identified by (a) the Discourse Skills Checklist (Bedrosian 1985), focusing on topic initiation (frequency and scope) and topic maintenance, and (b) the Conversation Analysis Coding Scheme (Larson and McKinley 2003), focusing on guiding elements such as "speaker roll" and "listener roll." Thus, the SCQ-T included three subscales: *speaker* (8 items, e.g., initiates conversation topic, expands and develops conversation, corrects communication flaws during conversation, ends conversation; $\alpha = 0.92$); *listener* (6 items, e.g., asks for clarification, responds to initiations, maintains conversation, takes turns; $\alpha = 0.89$); and *conversation functioning* (9 items, e.g., quality of participant's active participation via initiating/responding in peer conversation during various preschool activities like meals, creative activities, play; $\alpha = 0.89$). Items were rated on a 5-point Likert scale from 1 (*behavior does not exist*) to 5 (*behavior exists consistently*).

Uninvolved Head Preschool Educator's Ratings of Play and Conversation Skills To assess generalization of treatment efficacy, the uninvolved head educator of the preschool completed the SPQ regarding each participant's play stages and completed the SCQ regarding each participant's conversational skills as manifested in the child's everyday preschool behavior. See Appendix B for the educator-rated SPQ reliabilities. The educator-rated SCQ yielded the following internal consistency scores: $\alpha = 0.92$ for speaker, 0.87 for listener, and 0.87 for conversational functioning.

Informant Ratings to Examine Generalization of Trained Skills to Untrained Adaptive Behavior and Socialization as Manifested in Everyday Social Engagement in the Preschool

To assess generalization of trained skills to untrained areas of social functioning, we traced pre–post changes in the targeted preschoolers' adaptive behavior and socialization exhibited in the preschool, as reported by the uninvolved head educators (as full informants on generalization) and by the assigned therapists (as complementary informants). We used the Vineland Adaptive Behavior Scales-II: Teacher rating form (VABS-II; Sparrow et al. 2005) to assess participants' change in adaptive behavior (composite score) and in socialization level (using the standard score and its subdomains of interpersonal, play, coping). Lower VABS-II scores indicated more impaired social functioning.

Data Analysis

Within-group differences (T1/T2) and between-group differences (INTERACT/CONVERSE/PLAY/WAIT) were analyzed independently of each other within a generalized estimating equations (GEE) framework (SPSS V.24.0). Thus, for each participant, there were two repeats of each measurement: before (T1) and after (T2) intervention. The GEE is a generalization form of regression that allows an alternative assumption for the normal distribution, and a test for correlation between repeats, also known as the working matrix (Hardin and Hilbe 2013). We performed our test for group, time, and interaction effects.

The source of the time × group interaction was determined by using a post hoc pairwise comparison (LSD—least significant difference, subject to $p < 0.05$ significance level). We reported the marginal means' difference (*MD*), significance level (p), and confidence interval (CI) for our post hoc T2 results. All CIs were based on two-tailed 95% cut-off. We used the Benjamini–Hochberg correction (Benjamini and Hochberg 1995) to correct for false discovery rates in various research variables which represent the study's major experimental instruments for observed/reported peer engagement and generalization outcomes. We used this corrected version to verify our multiple test results. These corrected tests resulted in similar rejection decisions as the original multiple tests. The indicators included observations of social play and pragmatics; head educators' SPQ ratings on the most complex levels of social play (complementary–reciprocal) and of social pretend play (cooperative–complex); head educators' ratings on the SCQ speaker dimension; and head educators' Vineland composite score. The corrected decisions did not contradict the original decisions, thus we presented the multiple test results with higher confidence. To improve standard error estimation, we ran a three-level model to include the preschool effect at level 3. For this modeling approach, we used the HLM V.7.0 (Raudenbush et al. 2013).

Results

Preliminary analysis of between-group differences at T1 yielded non-significant findings. In line with study aims, and considering the non-significant T1 group differences, we report only main effects of time and interaction effects. Main group effect beyond time is seen in Tables 2, 3, 4 and 5.

The relevant tables also present the preschool level's intra class correlation (ICC) coefficients from the unconditional model. Commonly, a fraction above 0.05 is considered high enough to include the level in the model. The unconditional ICCs varied from as low as nearly zero to as high as 0.31 (see Tables 2, 3, 4 and 5). Thus, we ran all models as three levels and received the same results regarding the main effects (time and group) and the interactions (time × group). In these models, school varied randomly by the intercept only. Below we report the two-level GEE results as they are simpler and clearer to interpret.

Time Effects and Group × Time Interactions for Participants' Play Skills

Descriptive statistics and model results for play levels are presented in Table 2.

Table 2 Generalized estimating equations (GEE) results for observed and therapist-rated social and social pretend play

	Observation of free time in assigned group		Therapist ratings in preschool					
	Hierarchical Peer Play Scale		Social Play Questionnaire					
	Social play	Social pretend play	Social play			Social pretend play		
			Parallel	Parallel aware-simple	Complementary-reciprocal	Solitary	Coordinated- associative	Cooperative-complex
Preschool ICC	0.04	0.02	0.26	0.11	0.15	0.13	0.09	0.07
Effects (Wald)								
Group effect	4.21	10.84**	18.90***	7.25	26.12***	7.82*	16.49***	17.86***
Group diff.		P > I	I,C,W > P		P > I,C,W	P,C > W	I,C,P > W	I,C,P > W P > I,C,W
Time effect	14.00***	2.25	3.17	6.19**	27.96***	11.32***	0.21***	29.33***
Group × time	7.92*	8.89*	14.98**	14.72**	83.38***	10.30**	45.60***	83.87***
Before-after diff.	$I^{p=0.065}$ $P^{p=0.000}$	$P^{p=0.002}$	$I^{p=0.011}$ $P^{p=0.043}$	$C^{p=0.002}$ $P^{p=0.001}$	$I^{p=0.000}$ $P^{p=0.000}$	$I^{p=0.005}$ $P^{p=0.003}$	$I,P^{p=0.000}$	$I,P^{p=0.000}$ $W^{p=0.031}$
Group diff.–after	P > C,W	P > I,C,W	I,C,W > P	P > I,C,W	I,C,P > W I,C,W < P	I,C,P > W P > C	P > I,C,W I,C,P > W	P > I,C,W W < I,C,P
Marginal means (and standard error)								
Group								
INTERACT (I)	3.17 (0.13)	1.90 (0.24)	2.61 (0.11)	3.13 (0.14)	2.09 (0.18)	3.70 (0.17)	2.53 (0.15)	2.17 (0.14)
CONVERSE (C)	3.17 (0.21)	2.07 (0.27)	2.74 (0.18)	3.07 (0.16)	2.09 (0.11)	3.86 (0.14)	2.54 (0.15)	2.15 (0.15)
PLAY (P)	3.37 (0.09)	2.60 (0.24)	2.22 (0.12)	3.51 (0.14)	2.82 (0.13)	3.97 (0.16)	2.86 (0.12)	2.65 (0.15)
WAIT (W)	3.10 (0.10)	1.63 (0.18)	3.17 (0.18)	2.92 (0.23)	1.86 (0.18)	2.98 (0.32)	1.94 (0.19)	1.63 (0.18)
Time								
Before	3.02 (0.08)	1.90 (0.15)	2.77 (0.08)	3.03 (0.09)	1.83(0.09)	3.37 (0.15)	2.13 (0.11)	2.18 (0.10)
After	3.39 (0.09)	2.20 (0.16)	2.59 (0.09)	3.28 (0.10)	2.60 (0.12)	3.89 (0.10)	2.81 (0.10)	2.50 (0.10)
Group × time								
I—before	3.00 (0.16)	1.87 (0.30)	2.48 (0.12)	3.16 (0.09)	1.84 (0.15)	3.26 (0.27)	2.30 (0.17)	1.87 (0.14)
I—after	3.33 (0.09)	1.93 (0.29)	2.77 (0.12)	3.07 (0.15)	2.31 (0.22)	4.15 (0.19)	2.76 (0.16)	2.48 (0.16)
C—before	3.13 (0.25)	2.13 (0.35)	2.88 (0.16)	2.85 (0.19)	1.98 (0.18)	3.80 (0.25)	2.48 (0.24)	2.03 (0.22)
C—after	3.20 (0.19)	2.00 (0.30)	2.61 (0.12)	3.28 (0.16)	2.20 (0.13)	3.93 (0.06)	2.60 (0.14)	2.27 (0.13)
P—before	3.00 (0.07)	2.00 (0.23)	2.43 (0.19)	3.16 (0.17)	1.70 (0.15)	3.51 (0.29)	1.88 (0.16)	1.77 (0.16)
P—after	3.75 (0.18)	3.20 (0.36)	2.04 (0.13)	3.79 (0.16)	3.75 (0.18)	4.40 (0.13)	3.72 (0.16)	3.42 (0.17)
W—before	3.00 (0.09)	1.73 (0.26)	3.32 (0.19)	2.94 (0.21)	1.77 (0.15)	2.93 (0.37)	1.87 (0.20)	1.60 (0.18)
W—after	3.20 (0.17)	1.53 (0.31)	2.93 (0.29)	2.74 (0.29)	1.68 (0.22)	2.83 (0.25)	1.80 (0.19)	1.40 (0.14)

ICC for preschool-level intra class correlation coefficient from the unconditional three-level model

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

Table 3 Generalized estimating equations (GEE) results for observed and therapist-rated social conversation

	Observation of free time and snack time in assigned group				Therapist ratings in preschool		
	Pragmatics Rating Scale-Youth (PRS-Y)			Quality scale	Social Conversation Questionnaire		
	Pragmatic	Prosodic	Paralinguistic		Speaker	Listener	Functioning
Preschool ICC	0.09	0.03	0.17	0.31	0.13	0.04	0.12
Effects (Wald)							
Group effect	4.38	1.90	2.33	34.37***	3.92	3.40	6.53
Group diff.				C > I,W			
Time effect	6.09*	0.002	0.27	0.11	15.50***	25.98***	5.45*
Group × time	38.30***	13.82**	24.74***	87.29***	27.98***	26.18***	35.91***
Before-after diff.	$I^{p=0.023}$ $C^{p=0.000}$ $W^{p=0.001}$	$W^{p=0.003}$	$C^{p=0.000}$ $W^{p=0.001}$	$C^{p=0.000}$ $P^{p=0.033}$ $W^{p=0.050}$	$C^{p=0.000}$ $P^{p=0.023}$	$I^{p=0.045}$ $C^{p=0.000}$ $P^{p=0.002}$	$C^{p=0.000}$ $W^{p=0.050}$
Group diff.–after	W > I,C,P		W > C	C > I,P,W	I,C,P > W	W < C,P	C > I,P,W W < I,C,P
Marginal means (and standard error)							
Group							
INTERACT (I)	0.86 (0.08)	0.29 (0.05)	0.98 (0.09)	3.68 (0.17)	2.56 (0.14)	2.50 (0.15)	2.89 (0.20)
CONVERSE (C)	0.82 (0.09)	0.38 (0.06)	0.80 (0.09)	4.24 (0.06)	2.55 (0.16)	2.60 (0.13)	3.20 (0.21)
PLAY (P)	0.73 (0.04)	0.36 (0.04)	0.89 (0.06)	3.84 (0.08)	2.61 (0.15)	2.64 (0.14)	3.05 (0.19)
WAIT (W)	0.92 (0.08)	0.40 (0.11)	0.97 (0.12)	3.62 (0.11)	2.16 (0.19)	2.27 (0.16)	2.52 (0.19)
Time							
Before	0.88 (0.05)	0.35 (0.04)	0.93 (0.05)	3.86 (0.06)	2.30 (0.09)	2.28 (0.09)	2.79 (0.12)
After	0.78 (0.04	0.36 (0.04)	0.90 (0.05)	3.83 (0.08)	2.65 (0.09)	2.72 (0.08)	3.04 (0.10)
Group × time							
I—before	0.95 (0.11)	0.33 (0.05)	1.06 (0.09)	3.77 (0.14)	2.42 (0.18)	2.31 (0.19)	2.87 (0.26)
I—after	0.76 (0.07)	0.25 (0.05)	0.92 (0.11)	3.61 (0.22)	2.70 (0.19)	2.70 (0.17)	2.90 (0.17)
C—before	1.00 (0.12)	0.46 (0.08)	0.89 (0.09)	3.86 (0.08)	2.12 (0.18)	2.11 (0.16)	2.65 (0.26)
C—after	0.65 (0.07)	0.30 (0.09)	0.72 (0.09)	4.63 (0.05)	2.99 (0.16)	3.10 (0.14)	3.74 (0.19)
P—before	0.75 (0.05)	0.31 (0.05)	0.92 (0.07)	4.02 (0.08)	2.45 (0.15)	2.44 (0.16)	3.01 (0.22)
P—after	0.72 (0.08)	0.41 (0.06)	0.87 (0.09)	3.67 (0.14)	2.77 (0.17)	2.83 (0.16)	3.12 (0.21)
W—before	0.89 (0.09)	0.35 (0.11)	0.83 (0.14)	3.75 (0.13)	2.18 (0.21)	2.27 (0.18)	2.63 (0.22)
W—after	0.98 (0.08)	0.46 (0.11)	1.10 (0.10)	3.48 (0.13)	2.15 (0.18)	2.26 (0.16)	2.41 (0.18)

Higher PRS-Y scores indicated a more severe pragmatic deficit; ICC for preschool-level intra class correlation coefficient from the unconditional three-level model

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

Hierarchical-Developmental Play Stages (Peer Play Scale) in Small-Group Observations

Results of the GEE for coding of observed play behavior showed a main effect of time only for the hierarchical-developmental social play stage, with significantly higher T2 than T1 scores regardless of study group. A significant time × group interaction was found for both social play and social pretend play. The PLAY group demonstrated significant T1–T2 improvement on both social play ($p < 0.000$) and social pretend play ($p = 0.002$). Also, although not reaching statistical significance, the INTERACT group showed a tendency toward improvement ($p = 0.065$) on social play. At T2, the PLAY group scored significantly higher than the CONVERSE group ($MD = 0.55$, $p = 0.04$, CI [0.025, 1.075]) and significantly higher than the WAIT group ($MD = 0.55$, $p = 0.028$, CI [0.059, 1.041]) for their social play stage and significantly higher than all three other groups for social pretend play (INTERACT: $MD = 1.27$, $p = 0.006$, CI [0.37, 2.17]; CONVERSE: $MD = 1.20$, $p = 0.01$, CI [0.29, 2.12]; WAIT: $MD = 1.67$, $p = 0.000$, CI [0.74, 2.60]).

Therapist-Rated Play Stages (SPQ) in Preschool

A significant time effect was found in the expected direction (T1 < T2) for all therapist-rated SPQ play development categories except for the lowest level of social play (parallel play), beyond group. A significant time × group interaction

Table 4 Generalized estimating equations (GEE) results for uninvolved head educators' assessments of social/social pretend play and social conversation ratings in preschool

	Social Play Questionnaire						Social Conversation Questionnaire		
	Social play			Social pretend play			Speaker	Listener	Functioning
	Parallel	Parallel aware-simple	Complementary-reciprocal	Solitary	Coordinated- associative	Cooperative-complex			
Preschool ICC	0.36	0.09	0.15	0.21	0.19	0.10	0.09	0.12	0.10
Effects (Wald)									
Group diff.	4.08	4.84	5.99	4.72	11.55** P > C,W	23.74*** I,C,P > W P > C	6.27	6.40	5.01
Time effect	0.42	2.87	36.90***	2.34	24.17***	39.87***	21.44***	15.80***	9.06**
Group × time	12.89**	5.13	25.30***	7.72	21.01***	22.92***	42.11***	50.39***	21.60**
Before–after diff.	P$^{p=0.001}$	P$^{p=0.005}$	I$^{p=0.017}$ C$^{p=0.002}$ P$^{p=0.000}$	P$^{p=0.011}$	I$^{p=0.002}$ C$^{p=0.035}$ P$^{p=0.000}$	I$^{p=0.000}$ C$^{p=0.000}$ P$^{p=0.000}$	C$^{p=0.000}$ P$^{p=0.000}$	C$^{p=0.000}$ P$^{p=0.001}$	C$^{p=0.000}$
Group diff.–after	C > P	–	P > I,C,W I > W	I, P > W	P > I,C,W W < I,P	I,C,P > W P > I,C,W	W < I,C,P	P > W C > I,W	C,P > W
Marginal means (and standard error)									
Group									
INTERACT (I)	2.57 (0.17)	3.35 (0.10)	2.40 (0.17)	3.79 (0.20)	2.75 (0.20)	2.26 (0.19)	2.64 (0.16)	2.51 (0.16)	3.10 (0.17)
CONVERSE (C)	2.76 (0.15)	3.11 (0.09)	2.30 (0.16)	3.60 (0.22)	2.41 (0.13)	1.99 (0.14)	2.56 (0.13)	2.63 (0.15)	3.09 (0.17)
PLAY (P)	2.46 (0.15)	3.33 (0.17)	2.53 (0.15)	3.66 (0.17)	2.85 (0.23)	2.43 (0.12)	2.61 (0.17)	2.72 (0.17)	3.35 (0.19)
WAIT (W)	2.88 (0.18)	3.00 (0.16)	1.94 (0.19)	2.93 (0.34)	2.17 (0.23)	1.55 (0.14)	2.05 (0.20)	2.12 (0.19)	2.62 (0.26)
Time									
Before	2.69 (0.08)	3.11 (0.08)	1.85 (0.09)	3.34 (0.16)	2.17 (0.11)	1.64 (0.09)	2.23 (0.10)	2.29 (0.10)	2.86 (0.13)
After	2.64 (0.09)	3.29 (0.09)	2.73 (0.13)	3.65 (0.15)	2.92 (0.12)	2.47 (0.10)	2.70 (0.09)	2.69 (0.09)	3.23 (0.11)
Group × time									
I—before	2.46 (0.21)	3.39 (0.13)	2.11 (0.19)	3.66 (0.26)	2.49 (0.21)	1.98 (0.19)	2.62 (0.18)	2.50 (0.16)	3.02 (0.19)
I—after	2.69 (0.17)	3.29 (0.15)	2.68 (0.22)	3.93 (0.32)	2.99 (0.22)	2.53 (0.20)	2.66 (0.21)	2.52 (0.22)	3.19 (0.22)
C—before	2.77 (0.15)	3.03 (0.14)	2.04 (0.17)	3.60 (0.29)	2.19 (0.20)	1.65 (0.18)	2.09 (0.17)	2.17 (0.18)	2.59 (0.22)
C—after	2.74 (0.18)	3.18 (0.12)	2.55 (0.19)	3.60 (0.23)	2.62 (0.12)	2.32 (0.14)	3.03 (0.12)	3.10 (0.13)	3.59 (0.17)
P—before	2.69 (0.16)	3.08 (0.15)	1.61 (0.11)	3.16 (0.30)	2.02 (0.16)	1.62 (0.15)	2.23 (0.23)	2.41 (0.24)	3.19 (0.25)
P—after	2.23 (0.17)	3.56 (0.22)	3.44 (0.26)	4.16 (0.19)	3.69 (0.20)	3.25 (0.19)	2.99 (0.15)	3.02 (0.14)	3.51 (0.19)
W—before	2.79 (0.16)	2.96 (0.20)	1.80 (0.16)	3.08 (0.33)	2.17 (0.22)	1.48 (0.16)	2.06 (0.11)	2.19 (0.21)	2.70 (0.30)
W—after	3.02 (0.25)	3.03 (0.25)	1.99 (0.27)	2.67 (0.43)	2.08 (0.29)	1.52 (0.20)	2.05 (0.20)	2.05 (0.17)	2.56 (0.25)

ICC for preschool-level intra class correlation coefficient from the unconditional three-level model

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

emerged for all three social play levels and all three social pretend play levels. T1–T2 differences revealed the PLAY group's significant improvement on five levels ($p \leq 0.003$), all except the lowest social play level, where T1 scores were higher than T2 ($p = 0.043$). Participants in the INTERACT group likewise showed significant T1–T2 improvement on five SPQ developmental levels ($p \leq 0.01$), all except for the middle social play level. The CONVERSE group showed significant improvement only on the middle social play level ($p = 0.002$). By contrast, the WAIT group showed a significant decrease at T2 on the highest social pretend play level ($p = 0.031$).

Examination of group differences at T2 revealed that, for four of the six SPQ developmental levels, the WAIT group had significantly lower therapist-rated scores compared to all three intervention groups. They differed regarding the highest social play level of complementary-reciprocal play (INTERACT: $MD = -0.62$, $p = 0.04$, CI [−1.23, −0.016]; CONVERSE: $MD = -0.52$, $p = 0.039$, CI [−1.01, −0.025]; PLAY: $MD = -2.07$, $p = 0.000$, CI [−2.63, −1.51]).They

Table 5 Generalized estimating equations (GEE) results for therapists' and uninvolved educators' ratings on Vineland Adaptive Behavior Scale (VABS-II)

	Therapist ratings			Uninvolved educator ratings		
	Composite	Socialization	Play	Composite	Socialization	Play
Preschool ICC	0.18	0.25	0.20	0.00	0.03	0.11
Effects (Wald)						
Group effect	4.95	6.21	11.74***	5.23	2.90	6.84
Group diff.			3 > 4			
Time effect	1.87	2.47	4.40*	5.97**	1.26	3.63
Group × time	10.68**	10.39**	9.87*	10.93**	5.55	14.31***
Before-after diff.	P$^{p=0.000}$	P$^{p=0.006}$	P$^{p=0.000}$	I$^{p=0.021}$ P$^{p=0.004}$		P$^{p=0.000}$
Group diff.–after	I > C,W	P > W	P > I,C,W I > W	I,P > C,W		I,P > W
Marginal means (and standard error)						
Group						
INTERACT (I)	86.67 (2.46)	83.02 (2.84)	11.74 (0.51)	87.48 (2.39)	84.61 (2.22)	11.84 (0.41)
CONVERSE (C)	78.83 (2.57)	80.96 (3.08)	11.16 (0.58)	82.67 (2.27)	80.61 (2.34)	11.57 (0.49)
PLAY (P)	82.95 (2.19)	85.95 (1.99)	12.40 (0.34)	84.77 (1.62)	83.08 (1.56)	12.11 (0.34)
WAIT (W)	82.10 (2.35)	78.33 (2.42)	10.36 (0.51)	80.44 (2.23)	79.24 (2.91)	10.30 (0.60)
Time						
Before	81.86 (1.46)	81.05 (1.51)	11.11 (0.31)	82.52 (1.38)	81.24 (1.16)	11.19 (0.26)
After	83.42 (1.18)	83.08 (1.41)	11.71 (0.25)	85.16 (1.26)	82.52 (1.40)	11.72 (0.28)
Group × time						
I—before	85.79 (3.51)	83.32 (3.69)	11.75 (0.71)	85.20 (2.45)	85.36 (2.53)	11.84 (0.49)
I—after	87.60 (2.26)	83.00 (2.86)	11.73 (0.50)	89.47 (2.75)	84.06 (2.43)	11.86 (0.41)
C—before	78.86 (2.93)	79.93 (3.35)	10.73 (0.76)	83.50 (2.15)	79.92 (2.16)	11.50 (0.52)
C—after	78.80 (2.65)	82.00 (3.24)	11.60 (0.62)	80.79 (2.61)	81.26 (3.29)	11.50 (0.82)
P—before	80.20 (2.25)	82.60 (2.09)	11.70 (0.38)	81.96 (1.88)	80.65 (1.76)	11.20 (0.33)
P—after	85.70 (2.40)	89.30 (2.55)	13.10 (0.38)	87.28 (1.91)	85.16 (1.92)	12.88 (0.47)
W—before	83.26 (3.03)	79.40 (2.76)	10.46 (0.60)	79.80 (2.58)	79.61 (2.39)	10.39 (0.66)
W—after	80.93 (1.94)	77.26 (2.27)	10.26 (0.49)	81.23 (2.38)	79.11 (3.48)	10.28 (0.61)

ICC for preschool-level intra class correlation coefficient from the unconditional three-level model

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

also differed regarding all three social pretend play levels: *Level 1* (INTERACT: $MD = -1.32$, $p = 0.000$, CI [−1.96, −0.68]; CONVERSE: $MD = -1.09$, $p = 0.000$, CI [−1.62, −0.58]; PLAY: $MD = -1.56$, $p = 0.000$, CI [−2.13, −1.001]); *Level 2* (INTERACT: $MD = -0.95$, $p = 0.000$, CI [−1.45, −0.45]; CONVERSE: $MD = -0.79$, $p = 0.001$, CI [−1.27, −0.31]; PLAY: $MD = -1.92$, $p = 0.000$, CI [−2.40, −1.42]); and *Level 3* (INTERACT: $MD = -1.09$, $p = 0.000$, CI [−1.50, −0.67]; CONVERSE: $MD = -0.88$, $p = 0.000$, CI [−1.25, −0.51]; PLAY: $MD = -2.03$, $p = 0.000$, CI [−2.47, −1.58]).

Also, the PLAY group's SPQ scores at T2 were significantly higher than the CONVERSE group on the lowest social pretend play level ($MD = 0.60$, $p = 0.035$, CI [0.42, 1.16]) and were significantly higher than all three other groups on four of the play levels (middle and highest social play and social pretend play levels), as follows: social play *Level 2* (INTERACT: $MD = 0.72$, $p = 0.001$, CI [0.29, 1.15]; CONVERSE: $MD = 0.51$, $p = 0.027$, CI [0.06, 0.95]; WAIT: $MD = 1.05$, $p = 0.002$, CI [0.39, 1.71]); social play *Level 3* (INTERACT: $MD = 1.44$, $p = 0.000$, CI [0.88, 2.01]; CONVERSE: $MD = 1.55$, $p = 0.003$, CI [1.11, 1.99]; WAIT: $MD = 2.07$, $p = 0.000$, CI [1.51, 2.63]); social pretend play *Level 2* (INTERACT: $MD = 0.96$, $p = 0.000$, CI [0.51, 1.41]; CONVERSE: $MD = 1.13$, $p = 0.000$, CI [0.71, 1.55]; WAIT: $MD = 1.91$, $p = 0.000$, CI [1.42, 2.41]); and social pretend play *Level 3* (INTERACT: $MD = 0.93$, $p = 0.000$, CI [0.47, 1.40]; CONVERSE: $MD = 1.15$, $p = 0.000$, CI [0.72, 1.57]; WAIT: $MD = 2.03$, $p = 0.000$, CI [1.58, 2.47]). In contrast, for the lowest social play level, the PLAY group's score at T2 was significantly lower than all other groups. In sum, compared to all other groups after treatment, the WAIT group showed significant disadvantages, whereas the PLAY group showed more noticeable improvement in several complex social play and social pretend play stages.

Time Effects and Group × Time Interactions for Participants' Conversation Skills

Descriptive statistics and model results for conversation skills are presented in Table 3.

Interactive Social Conversation Skills (Pragmatics and Quality) Observed in Small Groups

Results of the GEE showed a main effect of time for the PRS-Y's pragmatic category, where deficit scores decreased from T1 to T2. A significant time × group interaction was found for all three PRS-Y categories (pragmatic, prosodic, and paralinguistic) and in the conversation quality score. Regarding T1–T2 differences for the WAIT group, all the PRS-Y deficit categories were significantly higher at T2 than at T1 ($p \leq 0.003$), and conversation quality was significantly lower at T2 ($p < 0.05$), meaning that on both conversation measures the waitlisted participants deteriorated from T1 to T2. The opposite pattern was shown by the CONVERSE group, with significantly lower pragmatic and paralinguistic deficits ($p = 0.000$) and higher conversation quality after treatment than before treatment ($p = 0.000$). A significant reduction in the pragmatic deficit at T2 was also observed for the INTERACT group ($p = 0.023$). By contrast, the PLAY group revealed a significant T1–T2 reduction in conversation quality ($p = 0.03$). Significant between-group differences on PRS-Y scores at T2 revealed the WAIT group's more severe pragmatic deficit compared to all intervention groups (INTERACT: $MD = 0.22$, $p = 0.041$, CI [0.009, 0.43]; CONVERSE: $MD = 0.34$, $p = 0.002$, CI [0.13, 0.55]; PLAY: $MD = 0.26$, $p = 0.009$, CI [0.06, 0.46]); and more severe paralinguistic deficit compared to the CONVERSE group ($MD = 0.39$, $p = 0.005$, CI [0.12, 0.66]); Lastly, between-group differences on conversation quality at T2 showed higher quality scores for the CONVERSE group than for any of the other groups (INTERACT: $MD = 1.02$, $p = 0.000$, CI [0.57, 1.47]; PLAY: $MD = 1.14$, $p = 0.000$, CI [0.85, 1.43]; WAIT: $MD = 0.96$, $p = 0.000$, CI [0.66, 1.25]).

Therapist-Rated Social Conversation Skills (SCQ) in Preschool

Results of the GEE revealed a main effect of time (T2 > T1) for all three therapist-rated SCQ categories (speaker, listener, conversational functioning). Likewise, a significant time × group interaction effect was found for all three categories. Therapists' ratings indicated significant T1–T2 improvement for the CONVERSE group in all categories ($p = 0.000$), for the PLAY group in the speaker and listener categories ($p = 0.023$, $p = 0.002$, respectively), and for the INTERACT group only in the listener category ($p = 0.045$). By contrast, therapists rated the WAIT group as showing a reduction from T1 to T2 in conversational functioning ($p = 0.05$). Between-group differences on the SCQ at T2 revealed that the WAIT group was rated significantly lower than all three intervention groups on the speaker category (INTERACT: $MD = -0.55$, $p = 0.040$, CI [−1.07, −0.02]; CONVERSE: $MD = -0.84$, $p = 0.001$, CI [−1.33, −0.35]; PLAY: $MD = -0.62$ $p = 0.014$, CI [−1.12, −0.12]) and on the functioning category (INTERACT: $MD = -0.49$, $p = 0.05$, CI [−0.98, −0.01]; CONVERSE: $MD = -0.98$, $p = 0.01$, CI [−1.85, −0.80]; PLAY: $MD = -0.70$, $p = 0.01$, CI [−1.25, −0.16]) and was rated significantly lower than the CONVERSE and PLAY groups on the listener category (CONVERSE: $MD = -0.83$, $p = 0.000$, CI [−1.26, −0.41]; PLAY: $MD = -0.57$, $p = 0.012$, CI [−1.001, −0.012]). In addition, the post-intervention rating for the CONVERSE group's conversational functioning was significantly higher than for all other groups (INTERACT: $MD = 0.84$, $p = 0.002$, CI [0.32, 1.35]; PLAY: $MD = 1.32$, $p = 0.000$, CI [0.80, 1.85]; WAIT: $MD = 0.62$, $p = 0.030$, CI [0.06, 1.17]).

Generalization: Uninvolved Educators' Reports on Extra-Group Play and Conversation

Descriptive statistics and model results for educator ratings are presented in Table 4. A significant time effect emerged-with higher scores at T2 than at T1—for three of the six educator-rated SPQ play development levels (the highest social play and middle and highest social pretend play levels) and for all three educator-rated SCQ conversation categories (speaker, listener, functioning). Regarding educator-rated play, a significant time × group interaction emerged for four SPQ levels (all but the middle social play level and lowest social pretend play level). Educators rated significant T1–T2 improvement for all intervention groups on three play levels: the highest social play (INTERACT: $p = 0.017$, CONVERSE: $p = 0.002$, PLAY: $p = 0.000$), the middle social pretend play (INTERACT: $p = 0.002$, CONVERSE: $p = 0.035$, PLAY: $p = 0.000$), and the highest social pretend play ($p = 0.000$ for all) levels. Only the PLAY group was rated by educators as showing a reduction from T1 to T2 on the lowest social play level ($p = 0.001$), alongside T1–T2 improvement in the lowest social pretend play level ($p = 0.01$). Main findings for between-group educator-rated SPQ differences at T2 revealed that the PLAY group surpassed all other groups on three ratings: social play *Level 3* (INTERACT: $MD = 0.76$, $p = 0.025$, CI [0.09, 1.43]; CONVERSE: $MD = 0.83$, $p = 0.006$, CI [0.25, 1.52]; WAIT: $MD = 1.45$, $p = 0.000$, CI [0.71, 2.19]); social pretend play *Level 2* (INTERACT: $MD = 0.70$, $p = 0.020$, CI [0.11, 1.29]; CONVERSE: $MD = 1.07$, $p = 0.000$, CI [0.61, 1.53]; WAIT: $MD = 1.61$, $p = 0.000$; CI [0.90, 2.31]); and social pretend play *Level 3* (INTERACT: $MD = 0.71$, $p = 0.01$, CI [0.17, 1.27]; CONVERSE: $MD = 0.92$, $p = 0.000$, CI [0.44,

1.40]; WAIT: $MD = 1.73$, $p = 0.000$, CI [1.18, 2.28]). Also, the WAIT group was lower than the three intervention groups on the highest social pretend play level (INTERACT: $MD = -1.01$, $p = 0.000$, CI [−1.57, −0.45]; CONVERSE: $MD = -0.80$, $p = 0.001$, CI [−1.29, −0.31]; PLAY: $MD = -1.73$, $p = 0.000$, CI [−2.28, −1.19]).

Regarding educator-rated conversation, the significant time × group interaction found for all SCQ categories revealed T1–T2 improvement for the CONVERSE group on all three categories (speaker, listener, functioning, $p = 0.000$) and for the PLAY group on the speaker and listener categories ($p \leq 0.001$). Between-group educator-rated SCQ differences at T2 revealed that the WAIT group showed significantly poorer speaker skills than all three intervention groups (INTERACT: $MD = -0.61$, $p = 0.039$, CI [−1.19, −0.03]; CONVERSE: $MD = -0.98$, $p = 0.000$, CI [−1.44, −0.51]; PLAY: $MD = -0.94$, $p = 0.000$, CI [−1.43, −0.44]) as well as poorer listener and functioning skills compared to the CONVERSE group (Listener: $MD = 1.04$, $p = 0.000$, CI [0.61, 1.48]; Functioning: $MD = -1.03$, $p = 0.001$, CI [−1.62, −0.43]) and the PLAY group (Listener: $MD = 0.97$, $p = 0.000$, CI [0.53, 1.41]; Functioning: $MD = -0.95$, $p = 0.003$, CI [−1.57, −0.32]). Lastly, educators rated the CONVERSE group as higher than the INTERACT group on the SCQ listener category at T2 ($MD = 0.58$, $p = 0.028$, CI [0.062, 1.09]).

Generalization to Participants' Adaptive Behavior

Descriptive statistics and model results for participants' adaptive behavior ratings are presented in Table 5. Results of the GEE for the VABS-II revealed a time effect (T2 > T1) for the educator-rated composite score and for the therapist-rated play score. A significant time × group effect emerged for the VABS-II composite and for the socialization play sub-domain according to both therapists' and educators' reports and for the VABS-II socialization categories according to only therapists' reports. A post hoc pairwise comparison to determine the interaction's source revealed the PLAY group's significant improvement from T1 to T2 based on the therapist-rated ($p = 0.000$) and educator-rated ($p = 0.004$) composite as well as on the therapist-rated and educator-rated play subdomain ($p = 0.000$, each) and the therapist-rated socialization scores ($p = 0.006$). Significant T1–T2 improvement also emerged for the INTERACT group on the educator-rated composite score ($p = 0.021$). T1–T2 changes in the interpersonal and coping subdomains were not significant for either reporter. Between-group differences at T2 showed a significantly higher therapist-rated composite score for the INTERACT group compared to the CONVERSE ($MD = 8.80$, $p = 0.012$, CI [1.95, 15.64]) and WAIT ($MD = 6.67$, $p = 0.026$, CI [0.81, 12.52]) groups. Therapist ratings at T2 for the PLAY group also showed higher socialization compared to the WAIT group ($MD = 12.03$, $p = 0.000$, CI [5.32, 18.74]) and higher play subdomain scores compared to all other groups (INTERACT: $MD = 1.37$, $p = 0.03$, CI [0.12, 2.61]; PLAY: $MD = 1.50$, $p = 0.039$, CI [0.073, 2.93]; WAIT: $MD = 2.83$, $p = 0.000$, CI [1.60, 4.06]). For the therapist-rated play subdomain at T2, the INTERACT group also surpassed the WAIT group ($MD = 1.47$, $p = 0.038$, CI [0.85, 2.85]). For the educator-rated composite score at T2, the PLAY and INTERACT groups exceeded both the CONVERSE group (PLAY: $MD = 6.48$, $p = 0.045$, CI [0.13, 12.82]; INTERACT: $MD = 8.67$, $p = 0.022$, CI [1.24, 16.01]) and the WAIT group (PLAY: $MD = 6.04$, $p = 0.048$, CI [0.048, 12.03]; INTERACT: $MD = 8.23$, $p = 0.024$, CI [1.09, 15.36]). On the educator-rated play subdomain, both PLAY and INTERACT groups outperformed the WAIT group (PLAY: $MD = 2.60$ $p = 0.001$, CI [1.09, 4.12]; INTERACT: $MD = 1.58$, $p = 0.032$, CI [0.14, 3.02]).

Discussion

The major aim of this novel RCT investigating the holistic ecological implementation of the manualized peer–peer PPSI intervention was to increase social engagement of preschoolers with HFASD during mixed interactions with their peers with HFASD and peers with TYP. Strengthening these children's ability to interact, play, and converse more effectively with their peers, especially at early ages in preschool, may mitigate the risk for poor social adaptation later in life (Hay et al. 2009; Vaughn et al. 2016). Using direct observations of spontaneous behavior and involved therapists' and uninvolved educators' reports at baseline and after intervention for three treatment groups versus a control group, we aimed to demarcate each treatment modality's efficacy beyond natural maturation, while examining modalities' generalization to preschool contexts outside the small-group setting and to untrained domains (adaptive and socialization capabilities).

Intervention Effects on Peer Engagement According to Observation and Therapist Report

Overall, the target children in all three PPSI groups showed improvement in various aspects of peer engagement after the intervention, whereas the target children in the control group did not progress over time and, on some measures, even deteriorated. As expected, assessments of preschoolers' spontaneous display of the targeted interactive conversational and play behaviors in non-PPSI social situations in the study group and in the preschool underscored that the PLAY intervention group improved most on indices of play, the CONVERSE group improved most on indices

of conversation, and the INTERACT group showed some improvement on both the play and conversation abilities.

More specifically, both direct observations as well as therapist reports indicated a notable rise in the developmental stage that the PLAY intervention group exhibited after treatment, for both social and social-pretend forms of play. This was the only group that revealed significant improvement in the social type of play—shifting from the lowest social play level of simple/aware parallel play before intervention ($M = 3.00$, range 2.86–3.14) toward the highest social play level of interactive complementary play after intervention ($M = 3.75$, range 3.38–4.11). Likewise, only the PLAY group improved so substantially on the social pretend play measure, from the lowest solitary pretend acts related to and coordinated with peers at T1 ($M = 2.00$, range 1.54–2.46) toward the simple associated pretend play up to cooperative pretend play at T2 ($M = 3.20$, range 2.50–3.90).

Group differences in play skills at T2 provide additional evidence of the PLAY group's unique improvement. On the hierarchical observation scale, this group outperformed the CONVERSE and WAIT groups on social play and outperformed all other groups on social pretend play. Therapist reports on the SPQ corroborated these observation results well, showing higher T2 than T1 ratings of the PLAY group on all play levels except for the most basic (parallel) level, as well as higher therapist ratings at T2 on the middle and highest developmental stages of social play and of social pretend play identified for the PLAY group compared to all other groups. Interestingly, although conversation was not its explicit focus, the PLAY group also showed improvement in speaking and listening skills over time (pre < post) according to therapists' questionnaires (SCQ) but not in the quality of their conversation, according to both the therapist-rated SCQ and the direct observations. This finding may indicate the importance of language for social kinds of play, especially during the preschool period; indeed, active play-talk is the most frequent type of conversation among preschoolers (Blum-Kulka et al. 2004). However, to improve the quality of such conversation during play, targeted training should probably be undertaken.

Indeed, preschoolers in the CONVERSE intervention showed consistent and notable T1–T2 improvement in their conversation quality, as observed in their smaller pragmatic deficit and more adequate paralinguistic skills (on the PRS-Y) after treatment. Higher conversation quality also denoted better attending skills during interactions, greater abilities to initiate conversation, more adequate responses to peers and, most importantly, a better ability to maintain, expand, and develop conversations with peers. This last skill is considered to be very challenging even for young children with TYP, and definitely for many children with ASD (e.g., Ying Sng et al. 2018). Therapists' reports confirmed this T1–T2 improvement in conversational qualities, and group differences at T2 showed higher gains on observed and reported conversation qualities for the CONVERSE group as compared to all the other groups.

Conversational and play skills are the building blocks of an efficient social interaction with peers (Coplan and Arbeau 2009). Indeed, the INTERACT intervention group showed improvement on both the play and conversation measures. Therapists reported significant improvement for this group on all but one of the six SPQ play development levels (all but the middle social play level). Likewise, children in the INTERACT group showed improvement in their pragmatic conversation deficit (PRS-Y) as well as better listening skills according to therapists' SCQ reports. In addition, on observed social play, pre–post scores showed a trend toward improvement but only barely reached statistical significance; thus, caution should be taken in considering this result. All in all, children's more general interactive skills such as prosocial capabilities, as promoted in the INTERACT intervention, led to better performance in some of their social conversational and play capabilities. However, according to both observations and therapists' reports, this group did not surpass either of the other intervention groups on play or conversation measures at T2. These findings may be related to our measurement procedure, which did not include an observation of children's spontaneous interactions with their peers during ordinary preschool activities (see "Limitations").

This study's most noticeable finding was the poor performance of the treatment-as-usual control group. Consistently over the various observed and reported measures, this WAIT group did not show improvement over the 6-month time period, and on several measures they even showed deterioration over time (e.g., on all PRS-Y categories, observed and reported conversational quality, highest reported social play level). Moreover, for many measures, they performed less well at T2 than all the other groups (e.g., on their highest social play level and all their social pretend play levels on the therapist-rated SPQ, observed PRS-Y pragmatics, and therapist-rated SCQ speaker and conversation functioning capacities). The poor within-group (pre–post) and between-group performance of the control group appears to suggest that peer interaction skills may not develop spontaneously in these preschoolers with HFASD, reinforcing the need to undertake specific interventions.

Generalization to Untrained Domains and Settings

The reports by each preschool's head educator (who was not involved in the treatment and was masked to the evaluated children's intervention group assignment), aiming to elucidate children's transfer of learned peer engagement skills, yielded several important findings. Overall, according to reports by these senior educators, all three intervention

groups showed significant T1–T2 increases in their play complexity (both in social play and social pretend play) as manifested during participation in various non-PPSI preschool activities, but children in the control group did not. Thus, learning and practicing of each of the PPSI's key domains helped children to develop important social engagement skills that were relevant for everyday play situations with their peers. Educator-rated group differences at T2 strengthened this conclusion, specifically for the middle and highest levels of social play, for which all three intervention groups outperformed the control group. Although the educators' evaluations of play skills were identical to those of the PLAY and WAIT groups' therapists, they varied somewhat from those of the INTERACT and CONVERSE therapists. Regarding conversation skill ratings, educators resembled therapists in denoting progress in all SCQ categories (speaker, listener, and functioning) for the CONVERSE intervention and some improvement in the speaker and listener categories for the PLAY group. This finding provides more support for the important role of language within play activities.

We further evaluated generalization of the PPSI by looking at skills that were not explicitly trained, such as general adaptive behaviors—as rated by therapists and uninvolved educators on the VABS-II. Both therapists and educators reported T1–T2 improvements in the PLAY group's adaptive scores, and educators also reported such improvement in the INTERACT group. This may suggest the important contribution made by interaction and social peer play to children's adaptive functioning during preschool. The therapists and educators also gave higher T1–T2 evaluations to the PLAY group on the VABS-II socialization score and play subdomain, indicating the major contribution of children's peer play capabilities to their socialization during the preschool years. Interestingly, the improvement that was observed and reported for the CONVERSE group in conversational skills was not reported by either the therapists or educators as contributing to better adaptive behavior or socialization. Perhaps these VABS-II domains were not sensitive enough to capture language improvement, calling for further study.

Limitations, Implications, and Future Research

Overall, the current outcomes were encouraging: the manualized PPSI peer engagement intervention led to increases in peer engagement among preschoolers with ASD. However, several study limitations should be considered. We directly observed free-play and snack time within a small group containing one or two children with HFASD and two age-mates with TYP. The latter were chosen based on their regular preschool educators' report stating their lack of significant social and behavior difficulties or a disability of any kind. Yet, our coding system included only the children with ASD because we were not permitted by the Ministry of Education to collect data on the children with TYP. Thus, the children with TYP may have exerted some influence on the behavior of the children with ASD that we did not capture. Hence, future studies would be well advised to look further at dyadic ASD–TYP influences.

Another limitation relates to instrumentation. A specialized measure to assess children's spontaneous interactions with peers during breaks in the school day would help to capture changes, mainly in the INTERACT intervention, but also to more objectively assess the generalization of children's outcomes, beyond evaluations by the educator and observations in the intervention group. Also, the SCQ and the SPQ were developed for the current study to evaluate children's improvement in interactive play and conversation skills; however, replication of those instruments in future studies is recommended to strengthen their validity. In addition, an important next step will be to examine children's progress following the PPSI interventions against normative standards, to shed light on our results' clinical significance. Indeed, the consistent progress exhibited by children in the treatment groups versus the consistent lack of pre–post progress (and even regression) shown by the control group speaks to the value of the former. Nevertheless, comparison of these improvements to normative standards may strengthen the clinical significance of the PPSI's results.

Moreover, collecting data throughout the course of the treatment could provide us with a closer look at children's progress in real time over the months of the intervention, to help make decisions about optimal intervention duration. Although the current study offered promising findings regarding the feasibility of the PPSI intervention to achieve meaningful change in young children with ASD, future attempts to unpack its clinically effective components and time requirements are important, considering the substantial resources needed for this intense 6-month treatment involving the preschool environment and peers as delivery agents. Specifically, further research should systematically compare treatment protocols that combine multiple treatment domains, simultaneously or sequentially. Also, future researchers would do well to look at the role of possible sex differences in understanding study outcomes, especially regarding conversational skills, in order to optimally adapt intervention goals to young boys' and girls' differing needs.

Overall, generalization and maintenance of study gains should be further explored, and direct exposure of school staff and parents to the intervention should further increase generalization. In a like manner, future empirical examination of individual differences in social behavior outcomes may be an important step toward personalizing the intervention to create a better fit with each child's needs. In that regard, it is important to note that we defined our ASD sample as "high-functioning" based solely on their cognitive

abilities (IQ > 70) as measured by the Mullen (1997) because the intervention required language capabilities above the intellectual disability cut-off. However, we are aware that IQ as a sole criterion does not address the many social or sensory challenges that may impact the way a child on the spectrum navigates social engagement with peers, thereby calling for future research to explore individual differences in response to treatment.

Conclusion

As seen in the current study, 3 h per week dedicated specifically to peer engagement in preschoolers with HFASD produced significant changes, suggesting the value of integrating peer-delivered training into early intervention models to reduce the risk of a poor social adaptation outcome (Hay et al. 2009; Vaughn et al. 2016). Informal positive reports of the therapists in their biweekly supervision sessions with the research team spoke to the social validity of the intervention. Furthermore, the peers with TYP seemed to enjoy taking part in the fun group activities. Based on our PPSI results that pinpointed the unique and crucial contribution of each of the three areas of peer engagement—play, interaction, and conversation—it appears that the incorporation of all three training modalities into early ASD intervention may create an optimal holistic peer program for supporting social and play outcomes in this age group.

As recommended above, future research should systematically unpack the PPSI's most effective components and domains; however, the current results already appear to offer some possible clinical implications. To increase feasibility for this time-intensive holistic peer intervention, it is advisable for implementation of the three treatment modalities (play, conversation, and interaction) to be individualized—tailoring them sequentially or perhaps simultaneously to fit each child's specific needs. Thus, based on expertise in all three key domains for effective peer engagement, the PPSI therapist can generate treatment priorities, foci, and sequencing for each child. For example, intervention may begin with the play curriculum if social play is identified as the most challenging domain in a specific preschooler's repertoire. Alternatively, a child's evaluation may lead the therapist to combine domains, for instance to emphasize the child's quality of peer talk (conversation) and prosocial capabilities (interaction) when working on play skills.

It is noteworthy that our first round of therapist training enabled 50 early interventionists to successfully integrate the PPSI into children's individual psychoeducational programs. Nevertheless, future implementation studies on integrative PPSI protocols can shed more light on those clinical–practical aspects in the real world. Taken together, strengthening the ability of young ASDs to interact more effectively with their peers may lead to a reduction in social withdrawal, thereby increasing quality of life and reducing the risk of poor long-term adjustment.

Acknowledgments This research was partially supported by the *Association of Children at Risk* given to the first author. The study was prepared based on doctoral dissertations of the second, third, and fourth authors under the guidance of the first author. Special thanks are extended to the children, preschools, and therapists who took part in this study. The authors would like to express their appreciation to Dee B. Ankonina for her editorial contribution.

Author Contributions The study was prepared based on doctoral dissertations of the second,third, and fourth authors under the guidance of the first author.

References

Adams, C. (2002). Practitioner review: The assessment of language pragmatics. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 43,* 973–987. https://doi.org/10.1111/1469-7610.00226.

Alvares, G. A., Bebbington, K., Cleary, D., Evans, K., Glasson, E. G., Maybery, M. T., et al. (2019). The misnomer of "high functioning autism:" Intelligence is an imprecise predictor of functional abilities at diagnosis. *Autism*. https://doi.org/10.1177/1362361319852831.

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

Bauminger, N. (2002). The facilitation of social-emotional understanding and social interaction in high functioning children with autism: Intervention outcomes. *Journal of Autism and Developmental Disorders, 32,* 283–298.

Bauminger, N. (2007a). Brief report: Group social-multimodal intervention for HFASD. *Journal of Autism and Developmental Disorders, 37,* 1605–1615. https://doi.org/10.1007/s10803-006-0246-3.

Bauminger, N. (2007b). Individual social-multimodal intervention for HFASD. *Journal of Autism and Developmental Disorders, 37,* 1593–1604. https://doi.org/10.1007/s10803-006-0245-4.

Bauminger-Zviely, N. (2013). *Social and academic abilities in children with high-functioning autism spectrum disorders*. New York: Guilford Press.

Bauminger-Zviely, N., Golan-Itshaky, A., & Tubul-Lavy, G. (2017). Speech acts during friends' and non-friends' spontaneous conversations in preschool dyads with high-functioning autism spectrum disorder versus typical development. *Journal of Autism and Developmental Disorders, 47,* 1380–1390. https://doi.org/10.1007/s10803-017-3064-x.

Bauminger-Zviely, N., Karin, E., Kimhi, Y., & Agam-Ben-Artzi, G. (2014). Spontaneous peer conversations in preschoolers with HFASD versus typical development. *Journal of Child Psychology and Psychiatry, 55,* 363–373. https://doi.org/10.1111/jcpp.12158.

Bauminger-Zviely, N., & Shefer, A. (2019). *The Autism Preschool Peer Interaction Observation Scale (APIOS): Naturalistic observation to assess adaptive and non-adaptive social-communicative skills during ongoing interaction with peers*. Paper presented at the Autism Conference, Ben-Gurion University, Israel.

Bedrosian, J. (1985). An approach to developing conversational competence. In D. Ripich & F. Spinelli (Eds.), *School discourse problems* (pp. 237–254). San Diego, CA: College Hill Press.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B, 57,* 289–300.

Blum-Kulka, S., Huck-Taglicht, D., & Avni, H. (2004). The social and discursive spectrum of peer talk. *Discourse Studies, 6,* 307–328. https://doi.org/10.1177/1461445604044291.

Blum-Kulka, S., & Snow, C. (2004). Introduction: The potential of peer talk. *Discourse Studies, 6,* 291–306. https://doi.org/10.1177/1461445604044290.

Bronfenbrenner, U. (1992). Ecological system theory. In R. Vasta (Ed.), *Annals of child development: Six theories of child development: Revised formulation and current issues* (pp. 178–249). London: Kingsly.

Campbell, S. B., Mahoney, A. S., Northrup, J., Moore, E. L., Leezenbaum, N. B., Brownell, C. A., et al. (2017). Developmental changes in pretend play from 22- to 34-months in younger siblings of children with autism spectrum disorder. *Journal of Abnormal Child Psychology*, *46*(3), 639–654.

Chang, Y. C., Shire, S. Y., Shih, W., Gelfand, C., & Kasari, C. (2016). Preschool deployment of evidence-based social communication intervention: JASPER in the classroom. *Journal of Autism and Developmental Disorders, 46*(6), 2211–2223. https://doi.org/10.1007/s10803-016-2752-2.

Conallen, K., & Reed, P. (2017). Children with autism spectrum disorder: Teaching conversation involving feelings about events. *Journal of Intellectual Disability Research, 61,* 279–291. https://doi.org/10.1111/jir.12339.

Coplan, R. J., & Arbeau, K. A. (2009). Peer interaction and play in early childhood. In K. H. Rubin, W. M. Bukowski, & B. Laursen (Eds.), *Handbook of peer interaction, relationships and groups* (pp. 143–161). New York: Guilford Press.

Dobson, D., & Dobson, K. (2009). *Evidence-based practice of cognitive behavioral therapy*. New York: Guilford Press.

Ganz, J. B., & Flores, M. (2008). Effects of the use of visual strategies in play groups for children with autism spectrum disorders and their peers. *Journal of Autism and Developmental Disorders, 38,* 926–940. https://doi.org/10.1007/s10803-007-0463-4.

Hardin, J. W., & Hilbe, J. M. (2013). *Generalized estimating equations* (2nd ed.). Boca Raton, FL: Chapman and Hall/CRC.

Hay, D. F., Caplan, M., & Nash, A. (2009). Beginning of peer relations. In K. H. Rubin, W. M. Bukowski, & B. Laursen (Eds.), *Handbook of peer interaction, relationships and groups* (pp. 121–142). New York: Guilford Press.

Hobson, J. A., Hobson, R. P., Malik, S., Bargiota, K., & Calo, S. (2013). The relation between social engagement and pretend play in autism. *British Journal of Developmental Psychology, 31,* 114–127. https://doi.org/10.1111/j.2044-835X.2012.02083.x.

Howes, C. (1980). Peer play scale as an index of complexity of peer interaction. *Developmental Psychology, 16,* 371–372. https://doi.org/10.1037/0012-1649.16.4.371.

Howes, C. (2011). Social play of children with adults and peers. In A. D. Pellegrini (Ed.), *The Oxford handbook of the development of play* (pp. 231–244). New York, NY: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195393002.013.0018.

Howes, C., & Matheson, C. (1992). Sequences in the development of competent play with peers: Social and social pretend play. *Developmental Psychology, 28,* 961–974. https://doi.org/10.1037/0012-1649.28.5.961.

Howes, C., Unger, O., & Matheson, C. (1992). *The collaborative construction of pretend: Social pretend play functions*. Albany, NY: State University of New York Press.

Jordan, R. (2003). Social play and autistic spectrum disorders: A perspective on theory, implications and educational approaches. *Autism, 7,* 347–360. https://doi.org/10.1177/1362361303007004002.

Kalyva, E., & Avramidis, E. (2005). Improving communication between children with autism and their peers through the 'circle of friends': A small-scale intervention study. *Journal of Applied Research in Intellectual Disabilities, 18*(3), 253–261. https://doi.org/10.1111/j.1468-3148.2005.00232.x.

Kamps, D., Thiemann-Bourque, K., Heitzman-Powell, L., Schwartz, I., Rosenberg, N., Mason, R., et al. (2015). A comprehensive peer network intervention to improve social communication of children with autism spectrum disorders: A randomized trial in kindergarten and first grade. *Journal of Autism and Developmental Disorders, 45,* 1809–1824. https://doi.org/10.1007/s10803-014-2340-2.

Katz, E., & Girolametto, L. (2015). Peer-mediated intervention for preschoolers with ASD: Effects on responses and initiations. *International Journal of Speech Language Pathology, 17,* 565–576. https://doi.org/10.3109/17549507.2015.1024166.

Koegel, L. K., Park, M. N., & Koegel, R. L. (2014). Using self-management to improve the reciprocal social conversation of children with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 44,* 1055–1063. https://doi.org/10.1007/s10803-013-1956-y.

Krasny, L., Williams, B. J., Provencal, S., & Ozonoff, S. (2003). Social skills interventions for the autism spectrum: Essential ingredients and a model curriculum. *Child and Adolescent Psychiatric Clinics of North America, 12*(1), 107–122.

Landa, R. J., Piven, J., Wzorek, M., Gayle, J., Chase, G., & Folstein, S. (1992). Social language use by parents of autistic individuals. *Psychological Medicine, 22,* 245–254. https://doi.org/10.1017/s0033291700032918.

Larson, V., & McKinley, N. (2003). *Communication solutions for older students: Assessment and intervention strategies*. Eau Claire, WI: Thinking.

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Jr., Leventhal, B. L., DiLavore, P. C., et al. (2000). The autism diagnostic observation schedule—Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders, 30,* 205–223. https://doi.org/10.1023/A:1005592401947.

Luckett, T., Bundy, A., & Roberts, J. (2007). Do behavioral approaches teach children with autism to play or are they pretending? *Autism, 11,* 365–388. https://doi.org/10.1177/1362361307078135.

Manning, M., & Wainwright, L. (2010). The role of high level play as a predictor of social functioning in autism. *Journal of Autism and Developmental Disorders, 40,* 523–533. https://doi.org/10.1007/s10803-009-0899-9.

Morrier, M. J., & Ziegler, S. M. T. (2018). I wanna play too: Factors related to changes in social behavior for children with and without autism spectrum disorder after implementation of a structured outdoor play curriculum. *Journal of Autism and Developmental Disorders, 48,* 2530–2541. https://doi.org/10.1007/s10803-018-3523-z.

Mullen, E. M. (1997). *Mullen scales of early learning*. Los Angeles, CA: Western Psychological Services.

Müller, E., Cannon, L. R., Kornblum, C., Clark, J., & Powers, M. (2016). Description and preliminary evaluation of a curriculum for teaching conversational skills to children with high-functioning autism and other social cognition challenges. *Language, Speech, and Hearing Services in Schools, 47,* 191–208. https://doi.org/10.1044/2016_LSHSS-15-0042.

Murdock, L. C., & Hobbs, J. Q. (2011). Picture me playing: Increasing pretend play dialogue of children with ASD. *Journal of Autism and Developmental Disorders, 41*(7), 870–878. https://doi.org/10.1007/s10803-010-1108-6.

Neff, E. R., Betz, A. M., Saini, V., & Henry, E. (2017). Using video modeling to teach siblings of children with autism how to prompt

and reinforce appropriate play. *Behavioral Interventions, 32,* 191–293. https://doi.org/10.1002/bin.1479.

Oppenheim-Leaf, M. L., Leaf, J. B., Dozier, C., Sheldon, J. B., & Sherman, J. A. (2012). Teaching typically developing children to promote social play with their siblings with autism. *Research in Autism Spectrum Disorders, 6,* 777–791. https://doi.org/10.1016/j.rasd.2011.10.010.

Paul, R., Landa, R., & Simmons, E. (2014). Assessing and treating communication. In J. McPartland, A. Klin, & F. Volkmar (Eds.), *Asperger syndrome: Assessing and treating high functioning autism spectrum disorder* (2nd ed., pp. 103–142). New York, NY: Guilford Press.

Paul, R., Orlovski, S. M., Marcinko, H. C., & Volkmar, F. (2009). Conversational behaviors in youth with high-functioning ASD and Asperger syndrome. *Journal of Autism and Developmental Disorders, 39,* 115–125. https://doi.org/10.1007/s10803-008-0607-1.

Reichow, B., & Volkmar, F. R. (2010). Social skills interventions for individuals with autism: Evaluation for evidence-based practices within a best evidence synthesis framework. *Journal of Autism and Developmental Disorders, 40,* 149–166. https://doi.org/10.1007/s10803-009-0842-0.

Rogers, S. J. (2000). Interventions that facilitate socialization in children with autism. *Journal of Autism and Developmental Disorders, 30,* 399–409. https://doi.org/10.1023/A:1005543321840.

Rubin, E., & Lennon, L. (2004). Challenges in social communication in Asperger syndrome and high-functioning autism. *Topics in Language Disorders, 24,* 271–285. https://doi.org/10.1097/00011363-200410000-00005.

Schuler, A. L., & Wolfberg, P. J. (2000). Promoting peer socialization and play: The art of scaffolding. In B. Prizant & A. Wetherby (Eds.), *Language issues in autism and pervasive developmental disorder: A transactional developmental perspective* (pp. 225–227). Baltimore, MD: Paul H. Brookes.

Shores, R. L. (1987). Overview of research on social interaction: A historical and personal perspective. *Behavioral Disorders, 12,* 233–241.

Sigman, M., & Ruskin, E. (1999). Continuity and change in the social competence of children with autism, Down syndrome, and developmental delays. *Monographs of the Society for Research in Child Development, 64*(1), 1–114. https://doi.org/10.1111/1540-5834/00002.

Skolits, G. J., & Richards, J. (2010). Designing and applying project fidelity assessment for a teacher implemented middle school instructional improvement pilot intervention. *The Canadian Journal of Program Evaluation, 24,* 133–156.

Sparrow, S. S., Cicchetti, D. V., & Balla, D. A. (2005). *Vineland adaptive behavior scales: Survey form* (2nd ed.). Circle Pines, MN: AGS Publishing.

Stefanatos, G., & Baron, I. S. (2011). The ontogenesis of language impairment in autism: A neuropsychological perspective. *Neuropsychology Review, 21,* 252–270. https://doi.org/10.1007/s11065-011-9178-6.

Stein, M. L., Berends, M., Fuchs, D., McMaster, K., Saenz, L., Yen, L., et al. (2008). Scaling up an early reading program: Relationships among teacher support, fidelity of implementation, and student performance cross different sites and years. *Educational Evaluation and Policy Analysis, 30,* 368–388. https://doi.org/10.3102/0162373708322738.

Strauss, K., Esposito, M., Polidori, G., Vicarid, S., Valerid, G., & Favaa, L. (2014). Facilitating play, peer engagement and social functioning in a peer group of young autistic children: Comparing highly structured and more flexible behavioral approaches. *Research in Autism Spectrum Disorders, 8,* 413–423. https://doi.org/10.1016/j.rasd.2014.01.002.

Thiemann, K. S., & Goldstein, H. (2004). Effects of peer training and written text cueing on social communication of school-age children with pervasive developmental disorder. *Journal of Speech, Language, and Hearing Research, 47,* 126–144.

Vaughn, B. E., Santos, A. J., Monteiro, L., Shin, N., Daniel, J. R., Krzysik, L., et al. (2016). Social engagement and adaptive functioning during early childhood: Identifying and distinguishing among subgroups differing with regard to social engagement. *Developmental Psychology, 9,* 1422–1434. https://doi.org/10.1037/dev0000142.

Watkins, L., O'Reilly, M., Kuhn, M., Gevarter, C., Lancioni, G. E., Sigafoos, J., et al. (2015). A review of peer-mediated social interaction interventions for students with autism in inclusive settings. *Journal of Autism and Developmental Disorders, 45,* 1070–1083. https://doi.org/10.1007/s10803-014-2264-x.

Watkins, L., O'Reilly, M., Kuhn, M., & Ledbetter-Cho, K. (2019). An interest-based intervention package to increase peer social interaction in young children with autism spectrum disorder. *Journal of Applied Behavior Analysis, 52,* 132–149. https://doi.org/10.1002//jaba.514.

Wilson, E. R., Wine, B., & Fitterer, K. (2017). An investigation of the MATRIX training approach to teach social play skills. *Behavioral Interventions, 32,* 278–284. https://doi.org/10.1002/bin.1473.

Wolfberg, P. J. (2016). Integrated play groups model: Supporting children with autism in essential play experiences with typical peers. In L. A. Reddy, T. Files-Hall, & C. Schaefer (Eds.), *Empirically based play interventions for children* (2nd ed., pp. 223–240). Washington, DC: American Psychological Association. https://doi.org/10.1037/14730-012.

Ying Sng, C., Carter, M., & Stephenson, J. (2018). A systematic review of the comparative pragmatic differences in conversational skills of individuals with autism. *Autism and Developmental Language Impairments, 3,* 1–24. https://doi.org/10.1177/2396941518803806.