Examining the Effects of Peer Mediation on the Social Skills of Students with Autism Spectrum Disorder as Compared to Their Peers

Stephanie L. Hart and Devender R. Banda
Texas Tech University

Abstract: This study examined the extent to which peer mediation increased, maintained, and generalized social initiations and responses of elementary students with autism spectrum disorder as compared to their typically-developing peers. A multiple-baseline across participants design was used to determine the impact of intervention. Results indicated that peer mediation increased, maintained, and generalized responses for all participants, and initiations remained stable. Additionally, all participants increased responses to levels within the range of responses demonstrated by their peer groups. Probes to assess the persistence of generalization effects were collected and analyzed.

Social skills difficulties in students with autism spectrum disorder (ASD) can particularly impact a student’s educational experience because they affect both academic and social development (Rao, Beidel, & Murray, 2008). In academic settings, students with social skills difficulties may struggle to make eye contact with teachers and peers, recognize nonverbal body language, participate in cooperative play with peers, successfully begin or sustain conversation, or develop lasting friendships (Watkins et al., 2015). Social skills difficulties may impact whole and small group instruction, cooperative learning, and unstructured play. For students with ASD, their struggle with social competence, particularly with peers, may seriously impact their success in classrooms (Camargo et al., 2014; Kamps et al., 2015; Rao et al., 2008).

Researchers can identify areas of social skills difficulties in participants by using systematic observation to compare performance levels of the participant to levels demonstrated by typically-developing peers in comparable settings (Flynn & Healy, 2012). Performance standards can be based on the average performance levels of a group or on the performance of particularly competent individuals (Van Houten, 1979). Normative peer data can be used to select socially-valid intervention goals (McGee, Feldman, & Morrier, 1997), set target levels for those goals (Tremblay, Strain, Hendrickson, & Shores, 1981), and continue intervention until those levels are reached (Ennis et al., 2013; Rogers, 2000). When a participant’s target behavior falls into the range of peer norms, then the intervention may be considered successful (Ennis et al., 2013; Van Houten, 1979).

Because peer norms vary from group to group (McGrath, Bosch, Sullivan, & Fuqua, 2003), several researchers used peer norms from the typically-developing peers in their studies rather than relying upon aggregate data on norms for typically developing children (Greenwood, Walker, Todd, & Hops, 1981; Gresham et al., 2010). Peer data in these studies were collected in the target setting prior to intervention (McGrath et al., 2003; Zanolli, Daggett, & Adams, 1996) or throughout the intervention (Gonzalez-Lopez & Kamps, 1997; Hundert & Houghton, 1992; Mundschenk & Sasso, 1995). Using norms collected from peers in the target environment ensures that participants are trained to interact at rates similar to those of their particular peers (McGrath et al.). Without normative data, researchers risk in-
creasing social interactions of students with ASD to abnormally high levels that cannot be sustained in a natural environment (McGee et al., 1997).

Peer mediation is a well-researched intervention that utilizes peers to improve social skills difficulties in students with ASD (Chan et al., 2009; Flynn & Healy, 2012; Maheady, Harper, & Mallette, 2001; Rogers, 2000; Stichter, Randolph, Gage, & Schmidt, 2007; Utley & Mortweet, 1997; Watkins et al., 2015). Peer mediation is a broad category of interventions where peers model or prompt selected social skills (Chan et al., 2009; Watkins et al., 2015). Peer-mediated interventions include peer modeling, peer-initiation training, peer tutoring (Maheady et al., 2001; Utley & Mortweet, 1997), and cooperative learning (Kamps et al., 2002). Peer modeling uses peers as exemplars of appropriate and competent behavior. Peer-initiation training teaches peers strategies to make social initiations to target students, which in turn provides peers with increased opportunities to respond (Maheady et al., 2001; Utley & Mortweet, 1997). Peer tutoring consists of pairs or groups of students working together on an instructional task (Harrewer & Dunlap, 2001). Cooperative learning is when students work in small groups to complete projects, solve problems, or accomplish common goals (Dugan et al., 1995; Kamps et al., 2002). Cooperative learning may be used to increase academic performance, promote team building, and enhance social interactions (Dugan et al., 1995). Researchers use a variety of procedures to train peers, including direct instruction to peers, brief peer trainings (Watkins et al., 2015), and training participants along with peers (Gonzalez-Lopez & Kamps, 1997; Kamps et al., 2002; Morrison, Kamps, Garcia, & Parker, 2001). In addition to increasing social skills (Di Salvo & Oswald, 2002; Kamps et al., 2015; Koegel, Kurikose, Singh, & Koegel, 2012), peer mediation has been used to increase academic skills (Kamps et al., 2002; Ledford & Wehby, 2015) and decrease challenging behaviors such as aggression, stereotypy, and self-injury (Kennedy & Shukla, 1995; McConnell, 2002).

Peer mediation is well-suited for use in general education classrooms because teachers can use peers to supplement traditional, teacher-led academic instruction (Chan et al., 2009; Utley & Mortweet, 1997; Watkins et al., 2015). Peer-mediated approaches impact overall student learning by increasing student on-task time and providing additional opportunities for students to respond, receive individualized help and encouragement, and receive positive feedback (Maheady et al., 2001; Utley & Mortweet, 1997). For students who struggle with social skills, peer-mediated teaching methods may supplement a teacher’s instruction and feedback to students, increase students’ social skills without separate social skills instruction, and enhance the social and academic skills of students without disabilities (Maheady et al., 2001). Peer mediation can be incorporated into daily activities in general education classrooms and provide students with opportunities to interact with a variety of peers to promote generalization (Chan et al., 2009; Watkins et al., 2015).

This study examined the extent to which peer mediation increased, maintained, and generalized social initiations and responses of elementary students with ASD as compared to their typically-developing peers. Probes to assess the persistence of generalization effects were collected and analyzed.

Method

Participants

Four participants with ASD were recruited for this study from public schools in a metropolitan area. Recruited participants were students a) in pre-kindergarten or first grade, b) diagnosed with ASD by the school district, c) placed in a general education setting for academic instruction, d) with social skills difficulties identified by their classroom teachers, e) without co-morbid physical disabilities, such as cerebral palsy, which would significantly impact use of fine motor skills, and f) without specific sensory disabilities, such as blindness, deafness, deafblindness, or hearing loss, which would significantly impact use of printed materials or ability to converse with peers.

Participants who demonstrated difficulties with social skills were nominated by their teachers and screened using language and social skills assessments. To assess social skills, participants were administered the Social Skills Rating System (SSRS) Teacher Form (Gresham & Elliott, 1990). The SSRS is a Likert-type standardized rating form that measures social
skills, problem behaviors, and academic competence in school settings. The SSRS can be used to identify acquisition and performance difficulties in social behaviors. Participants’ classroom teachers completed the Elementary Level Form for students aged 5 years and above and the Preschool Level Form for students younger than 4 years, 11 months. To establish pre-study language abilities, participants were screened using two norm-referenced assessments of receptive and expressive vocabulary skills: the Receptive One-Word Vocabulary Test, Fourth edition (ROWPVT-4; Martin & Brownell, 2010) and the Expressive One-Word Picture Vocabulary Test, Third edition (EOWPVT-3; Martin & Brownell, 2000). See Table 1 for a summary of results from pre-study assessments.

John. John, a pre-kindergarten student, was 5 years, 3 months old at the start of the study. While his classmates were on average 3 ½ to 4 ½ years old, his parents requested that he repeat pre-kindergarten. Conclusions from the social skills screening were that John demonstrated below average academic skills, high levels of problem behaviors (fidgets, easily distracted, interrupts conversations, disturbs ongoing activities, acts impulsively, and has trouble transitioning between activities) and significantly below average social skills (rarely initiates conversation with peers, rarely initiates activities, has trouble transitioning activities, interrupts conversations, acts impulsively, and has trouble transitioning between activities). John’s academic competence was high, but he had difficulty with social skills.

Brad. Brad was a first-grade student in a self-contained, special education classroom. He attended a first grade, general education classroom for academic activities for one hour per day prior to the study. Brad was 7 years, 2 months old at the start of the study. While his classmates were on average 5 years, 3 months old, his parents requested that he repeat first grade. Conclusions from the social skills screening were that Brad demonstrated average academic skills, moderate levels of problem behaviors (fidgets, easily distracted, likes to be alone, interrupts the conversation, disturbs ongoing activities, acts impulsively, and has trouble transitioning between activities) and below average social skills (rarely initiates conversation with peers, rarely initiates activities, has trouble transitioning activities, interrupts conversations, acts impulsively, and has trouble transitioning between activities). Brad’s academic competence was high, but he had difficulty with social skills.

Alton. Alton, a pre-kindergarten student, was 4 years, 10 months old at the start of the study. Conclusions from the social skills screening were that Alton demonstrated high levels of social skills, average academic skills, and below average problem behaviors (fidgets, easily distracted, interrupts conversations, disturbs ongoing activities, acts impulsively, and has trouble transitioning between activities). Alton’s social skills were significantly below average, as he rarely invited others to join activities, rarely initiated activities, and had trouble transitioning activities.

Table 1

<table>
<thead>
<tr>
<th>Chronological Age</th>
<th>Receptive Vocabulary Age Equivalent</th>
<th>Expressive Vocabulary Age Equivalent</th>
<th>Academic Competence</th>
<th>Problem Behaviors</th>
<th>Social Skills</th>
</tr>
</thead>
<tbody>
<tr>
<td>John</td>
<td>5 y, 3 mos</td>
<td>3 y, 7 mos (9th percentile)</td>
<td>2 y, 8 mos (1st percentile)</td>
<td>88 (21st percentile)</td>
<td>110 (75th percentile)</td>
</tr>
<tr>
<td>Brad</td>
<td>7 y, 2 mos</td>
<td>3 y, 6 mos (1st percentile)</td>
<td>3 y, 6 mos (1st percentile)</td>
<td>103 (58th percentile)</td>
<td>102 (55th percentile)</td>
</tr>
<tr>
<td>Alton</td>
<td>4 y, 10 mos</td>
<td>3 y, 7 mos (10th percentile)</td>
<td>3 y, 5 mos (10th percentile)</td>
<td>N/A</td>
<td>120 (91st percentile)</td>
</tr>
<tr>
<td>Evan</td>
<td>6 y, 11 mos</td>
<td>5 y, 11 mos (32nd percentile)</td>
<td>5 y, 1 mo (13th percentile)</td>
<td>103 (58th percentile)</td>
<td>90 (25th percentile)</td>
</tr>
</tbody>
</table>

1 The minimum age for the Elementary Level Form of the SSRS was 4 years and 11 months. Therefore, Alton was scored using the Preschool Level Form, which did not assess academic competence.
problem behaviors (temper tantrums, fidgets or moves excessively, disturbs ongoing activities, disobeys rules or requests, and shows anxiety about being with a group of children) and below average social skills (rarely introduces self to others, invites others to join activities, cooperates with peers, waits turn in games or other activities, follows rules when playing games with others, or controls temper in conflict situations with peers). The minimum age for the Elementary Level Form of the SSRS was 4 years and 11 months. Therefore, Alton was scored using the Preschool Level Form, which did not assess academic competence.

Evan. Evan, a first-grade student, was 6 years, 11 months old at the start of the study. Conclusions from the social skills screening were that Evan demonstrated average academic skills, low levels of problem behaviors, and significantly below average social skills (rarely invites others to join in activities, initiates conversation with peers, volunteers to help peers, joins ongoing activity or group, cooperates with peers, or makes friends easily).

Typically-developing peers. For each participant, three typically-developing peers from the participant’s classroom were nominated by the student’s classroom teacher. Selected peers were students (a) in pre-kindergarten and first grades, (b) placed in the same general education class as one of the participants with ASD, (c) who demonstrated strong peer-to-peer social skills as observed by his/her classroom teacher, and (d) had no documented disabilities that required substantial intervention or services. Participants were approximately the same age as their peers except for John, who was repeating pre-kindergarten and therefore older than his peers. John, Brad, and Alton each had three peers in their group. John’s group had three boys, and Brad and Alton’s groups each had one boy and two girls. Evan had four peers (two boys and two girls) in his group. On days that Evan’s group had all four peers present, three peers were randomly chosen to participate in the group for that day.

Settings

The study was conducted in each participant’s general education classroom within the context of “center-time” activities. “Center-time” occurs when a class with 20–25 students is divided into small groups to complete academic tasks. The groups rotate through activities approximately every 15 minutes, either by physically moving to another area or by getting new materials. During center-time activities, the teacher may provide direct instruction for one group while simultaneously monitoring the progress and behavior of the groups. Although tasks used in center-time activities vary by teacher and grade level, they commonly include math, reading, and writing skills. The activities are completed independently or with other group members rather than with teacher assistance. In this study, each “target group” contained one student with ASD and three typically-developing peers from the same class. Because intervention took place during “center-time” activities occurring for all students in the class, students in the “target group” were not removed from their class and did not miss academic activities.

Materials

Specific center-time activities were selected for use in the study. Criteria for selected activities were as follows: (a) utilized shared materials rather than individual sets of materials, (b) allowed for social interaction among group members, and (c) had an academic focus. Most materials were educational games that encouraged social interactions by requiring players to take turns, follow game rules (e.g., players help one another count spaces on each move to ensure fair play), contribute pieces to a common board (e.g., word puzzles), or complete an academic task in order to move ahead in the game (e.g., reading games). Materials used in the pre-kindergarten class were counting mats (where students matched small objects to numbers on a mat), counting games (Hi! Ho! Cherry-O™), matching puzzles (What Goes Together?™ and Word Building Puzzles™), rhyming puzzles (Match It!™), play dough with carving tools, markers, paint daubers, and paper. The play dough, markers, and paint daubers were used for pre-writing tasks, such as tracing the first letter in a student’s name. Materials used in the first-grade classrooms were rhyming games (Chip-O™), phonics games (Max’s At-
tic™), sentence-building games (Silly Sentences™), and counting games (Chutes and Ladders™). When possible, materials were alternated so that the same materials were not used for consecutive sessions with the same group.

**Dependent Variables**

Two variables, initiations and responses, were recorded for participants and typically-developing peers. When making initiations and responses, participants and peers were not required to make eye contact, physically turn one’s body towards another student, or use complete sentences. Initiations were defined as the first verbal utterance in a social interaction and included greetings, asking questions, making requests, commenting, and saying a peer’s name (Goldstein, Kaczmarek, Pennington, & Shafer, 1992; Morrison et al., 2001). Initiations could reference any topic and were not limited to referencing the activity materials used in the study. Responses were defined as an on-topic, verbal reply to an initiation that occurred within 10 s of the initiation. Examples were answering questions, responding to requests, and commenting (Goldstein et al., 1992; Morrison et al., 2001). Responses were considered to be on-topic when they addressed the same topic as the initiation. If the initiation was a question about the game, then examples of on-topic responses include “It’s my turn” and “This is my favorite game” and non-examples include “Do you like Spider Man?” and “I like pizza.” On-topic comments were coded as responses, and off-topic comments were considered attempts to begin a new topic and were therefore coded as initiations. When no student in the group had made any verbal utterances for 10 s, the next verbal utterance was also recorded as an initiation. For both initiations and responses, if a student repeated what another student had just said or repeated his or her own comment, then the utterance was not coded.

The dependent variables were recorded for the participant and three peers, for a total of four students per group. These numbers were selected so that the group had enough members to generate conversation but not so many members that members could hold multiple, simultaneous conversations. On days when several peers were absent, resulting in a group with fewer than three students (including the participant), data were not collected.

For the purposes of collecting data, the students in each group were assigned a seat and a number (1-4) to be used across phases. Students were told to sit in their assigned seat for each session but were not informed that they were given a number. This procedure provided anonymity for the students by concealing from the second coder which students were typically-developing peers and which students had ASD. During coding, each student’s number was used to represent the social interactions of that student. Data were recorded using a frequency count to determine the number of initiations and responses for each student during a 10-minute period. When an initiation-response exchange occurred, the number of each student speaking was recorded on coding forms in this order: the initiator’s number first, the respondent’s number second, then the number of any student producing additional responses.

The researcher, the first author in the study, conducted the intervention and collected all data. A trained behavior analyst, who was not associated with the study, served as the second coder to collect interobserver agreement and procedural integrity data for approximately 30% of the sessions. All sessions were recorded using a video camcorder. An omnidirectional microphone was used to amplify the voices of students in the target group and helped to distinguish their voices from the general noise in the classroom. All data were coded from videotaped recordings. No data were coded live.

**Independent Variable**

The intervention was a 5-minute training session immediately followed by a 10-minute observation session, both occurring during 15-minute center time activities. Training sessions included aspects of direct instruction, including modeling, prompting, error correction, praise, and guided practice. During training, participants and peers were trained together to make initiations and responses. Participants and peers were treated equally during training, and participants were not isolated or identified to the group as “target students.” Training sessions are described in detail in the Procedures section.
Experimental Design

A single-subject multiple baseline across participants design was used to examine the effect of peer mediation across intervention, maintenance, and generalization phases. Maintenance and generalization were measured simultaneously to isolate problems with sustainability versus transfer (Koegel & Rincover, 1977). Graphs were visually analyzed to detect patterns in the data. Within each phase, data were analyzed by level, trend, and variability. Between phases, data were analyzed by immediacy of effect and overlap. Experimental control was evaluated based on the extent to which behavioral change was replicated across participants.

Procedure

The university IRB approved the study. Written consent for screening and participation was obtained from parents, and verbal assent was obtained from participants and peers. After obtaining consent, the researcher conducted the study in four phases: baseline, intervention, maintenance, and generalization. Students who exhibited problem behaviors in any phase were redirected back to the activity.

Baseline. Baseline observations took place during center time activities in each participant’s general education classroom. The center time activities lasted approximately 15 minutes, and baseline data were collected during the last 10 minutes of each session. During baseline, the classroom teacher instructed all students in a class to complete a small-group academic activity (i.e., center time) using teacher-selected materials. The groups sat at tables or on the floor (the target group sat at a table) and continued the activity until the classroom teacher signaled that that the activity was over. No instructions or prompts were provided other than to verbally remind students in the target group to stay at their table.

Intervention (training and observation). The intervention was a 5-minute training session immediately followed by a 10-minute observation session, both occurring during 15-minute center time activities. To begin a training session, the researcher sat with the target group of students at a table during center time and announced that it was time to practice talking to friends and answering friends’ questions. The researcher used either center-time activity materials that fit the above criteria (materials that promote social interaction) or developmentally appropriate games and arranged materials in the middle of the table or floor.

First, the researcher modeled initiations in the form of comments and questions. To model initiations, she made a comment about the activity (e.g., “Oh, I love this game!” or “It is your turn to go first this time.”) and said, “To start talking, I can make a comment about the activity.” To model initiative questions, she asked a student in the group a question about the activity (e.g., Can you please give me the ___? or “Do you need help?”) and said, “To start talking, I can ask a question.” When the researcher modeled a question, she directed the question to a specific student by using his/her name. The student was allowed to answer the question but was not prompted to answer if he/she failed to do so. Then the researcher asked each student to practice the skill by making a comment to another student or asking another student a question. After making the request, she paused to wait for the student to make a comment or ask a question. If the student did not demonstrate the skill within 5 seconds of the request, she verbally prompted the student. First, she used the student’s name and repeated the request. If the student did not make a comment or ask a question within 5 seconds, she asked another student to model an appropriate initiation as a form of error correction. All student attempts were verbally praised.

Next, the researcher modeled responses in the form of comments and direct responses. To model response comments, she made a comment about a specific aspect of the activity that was being discussed (e.g., “Oh, that is my favorite game, too!” or “You just moved ahead of me.”) and said, “To answer, I can make a comment about what you just said.” To model direct responses, she asked a student in the group to ask a question about the activity (e.g., Can you please give me the ___? or “Do you need help?”), modeled how to answer the student’s question, and said, “To keep talking, I can answer the question.” Then the researcher asked each student to practice the skill by making a response comment or responding to a question. She asked a student in
the group to ask a question about the activity and paused to wait for each student to make a response comment or answer the question. All students responded to the same question. If a student did not demonstrate the skill within 5 seconds of the request, she verbally prompted the student. First, she used the student’s name and repeated the request. If the student did not make a response comment or answer the question within 5 seconds, she asked another student to model an appropriate initiation as a form of error correction. All student attempts were verbally praised.

Observation sessions were used to observe how often participants or peers demonstrated the social skills taught during training. To begin a 10-minute observation session, the researcher thanked the students in the target group for participating in the training session, moved to a remote part of the classroom, and did not provide any prompts other than redirection. Students continued to use the same center-time activity or materials used during training. All students in the classroom (including students in the target group) continued to work until the classroom teacher signaled that the activity was over.

Maintenance and Generalization. Maintenance and generalization phases were measured simultaneously to isolate problems with sustainability versus transfer (Koegel & Rincover, 1977). Rather than having a maintenance phase and then a generalization phase, maintenance and generalization were probed several times each on different days in a random order. Maintenance and generalization were each probed for a minimum of two data points per participant over a period of two weeks after intervention ended. For maintenance sessions, target groups used training materials. Generalization was assessed using novel materials (i.e., academic materials not previously introduced in training sessions or used in the classroom). Maintenance and generalization data were collected during the last 10 minutes of each 15-minute center time activity and followed baseline procedures (small groups independently completed academic activities; no instructions or prompts were provided other than to verbally remind students in the target group to stay at their table).

Interobserver Agreement

The second coder, a board-certified Behavior Analyst, was trained to code interobserver agreement data from videotaped recordings for 30% of sessions. During training, the coder and the researcher independently viewed and coded videos, discussed discrepancies, and reviewed operational definitions until IOA was above 80% for three consecutive videos. Using a total agreement approach, interobserver agreement was calculated for each variable by dividing the smaller total by the larger total in each session, multiplying the amount by 100%, and averaging session totals. For initiations, each participant had 100% interobserver agreement. For responses, interobserver agreement was as follows: John, 77% (range 60%–94%); Brad, 82% (range 67%–100%); Alton, 80% (range 68%–91%); and Evan, 79% (range 69%–100%). While 80% is often used as the standard convention for observational recording, a mean IOA as low as 75% may be acceptable during “the simultaneous measurement of multiple behaviors by several subjects in a complex environment” (Cooper, Heron, & Heward, 2007). In two sessions where agreement fell below 60%, data points from those sessions were dropped from the study.

Procedural Integrity

The second coder also collected procedural integrity data for 30% of intervention sessions for each group. Procedural integrity was measured using a 10-item checklist of intervention procedures (e.g., “Models asking a question” and “Verbally praises student attempts”). Procedural integrity for each participant was as follows: John, 90% (all scores of 90%); Brad, 100% (all scores of 100%); Alton, 100% (all scores of 100%); and Evan, 100% (all scores of 100%). Due to scheduling conflicts, the researcher often omitted the first item on the checklist, “Sets up materials,” during John’s sessions because center time activities were already in progress when she arrived. Although this did not prevent the researcher from videotaping the full session (5 minutes for intervention and 10 minutes for observation), it did lower John’s procedural integrity average to 90%. Overall, the procedural integrity for the study was 97.8%.
Social Validity

Social validity was assessed in three areas: social validation of behavior goals, social appropriateness of treatment procedures, and social importance of results (Ennis et al., 2013; Wolf, 1978). Social validation of behavior goals was established using teacher responses on the SSRS (Gresham & Elliott, 1990) to evaluate participants’ skills in relation to typically-developing peers before intervention. Social appropriateness of treatment procedures was assessed by having classroom teachers and typically-developing peers respond to treatment acceptability questionnaires. Teachers filled out a modified version of the Intervention Rating Profile-15 (IRP-15; Martens, Witt, Elliott, & Darveaux, 1985), a 15-item questionnaire with a 6-point Likert-type scale that addressed treatment acceptability, effectiveness, and fairness. Results from the modified IRP-15 were calculated by summing item responses in each rating category (i.e., agree or strongly agree), dividing the sum in each rating category by the total number of items (15), and then multiplying by 100. Peers were individually read aloud a modified version of the Children’s Intervention Rating Profile (CIRP; Witt & Elliott, 1985) that contained open-ended questions (e.g., “What did you like about Games Group?”) and asked to respond verbally while the researcher recorded their responses. Each peer was told which student was the “target student” in his/her group so the peer could evaluate the target student’s progress (i.e., “Did Games Group make your friend a better talker?”) Results from the modified CIRP were reported anecdotally. Lastly, social importance of results was evaluated by comparing participant progress to local peer norms (Van Houten, 1979).

Results

Effects on Participants

Figure 1 shows the results of initiations and responses, maintenance, and generalization across participants. Results for each participant are described in the following sections.

John. John demonstrated averages of 0.7 initiations (range 0–1) during baseline, 2.2 initiations (range 1–5) during intervention, 0.5 initiations (range 0–1) during maintenance and no initiations during generalization. The data demonstrate an increase in John’s initiations from baseline, with some overlapping data points. His average level of initiations decreased from intervention to maintenance with some overlapping data points and fell to zero during generalization. For responses, John demonstrated averages of 2.6 responses (range 2–3) during baseline, 9.5 responses (range 7–13) during intervention, 18.6 responses (range 18–20) during maintenance, and 31.3 responses (range 21–42) during generalization. His responses increased from baseline to intervention with no overlapping data points, from intervention to maintenance with no overlapping data points, and from maintenance to generalization with no overlapping data points, even though maintenance and generalization were collected simultaneously.

Brad. Brad demonstrated no initiations during baseline and an average of 1.8 initiations (range 0–3) during intervention, 1.0 initiations (all scores of 1) during maintenance, and 1.5 initiations (range 1–2) during generalization. This indicates an increase in initiations from baseline to intervention with no overlapping data points and from intervention to maintenance with some overlapping data points. Initiations remained stable from intervention to generalization. For responses, Brad demonstrated averages of 3.0 responses (range 1–5) during baseline, 16.8 responses (range 7–24) during intervention, 26.0 responses (range 23–30) during maintenance, and 19.5 responses (range 19–20) during generalization. This indicates an increase in responses from baseline to intervention with no overlapping data points and a slight increase from intervention to maintenance and generalization with some overlapping data points.

Alton. Alton demonstrated averages of 1.1 initiations (range 0–3) during baseline, 0.4 initiations (range 0–2) during intervention, 0.5 initiations (range 0–1) during maintenance, and no initiations during generalization. This indicates a decrease in initiations from baseline to intervention with some overlapping data points. Average levels of initiations remained stable from intervention to maintenance and fell to zero during generalization. For responses, Alton demonstrated averages of 11.0 responses (range 3–23) during baseline, 18.8 responses (range 7–38) during
intervention, 28.0 responses (range 25–31) during maintenance, and 32.0 responses (range 29–35) during generalization. The data indicate an increase in responses from baseline to intervention and from intervention to maintenance and generalization with some overlapping data points. There is a slight downward trend within maintenance and generalization phases.

Evan. Evan demonstrated averages of 0.3 initiations (range 0–1) during baseline, 0.3 initiations (range 0–1) during intervention, no initiations during maintenance, and an average of 0.7 initiations (range 0–1) during generalization. The data demonstrate that Evan’s level of initiations remained stable throughout all phases. For responses, Evan demonstrated averages of 9.5 responses (range 3–18) during baseline, 33.1 responses (range 21–49) during intervention, 44.3 responses (range 43–46) during maintenance, and 58.0 responses (range 50–66) during generalization. His responses increased from baseline to intervention with no overlapping data points, from intervention to maintenance with some overlapping data points, and from maintenance to generalization with no overlapping data points, even though
maintenance and generalization were collected simultaneously.

**Effects on Participants as Compared to Peers**

To determine the effectiveness of the intervention, we compared average levels of initiations and responses between students with ASD and their peer groups across phases. To create a peer group average or local peer norm, we averaged the initiations in each peer group and the responses in each peer group by phase. Participant levels were not included in peer phase averages. Initiations for participants and peer groups are shown in Figure 2, and responses for participants and peer groups are shown in Figure 3. Overall, the level and trend of participant data often cor-

![Figure 2. Initiations for participants and peer groups across phases.](image-url)
responded with the level and trend of peer group data, particularly once intervention began.

John and his peers. John’s peers gradually decreased their levels of initiations from baseline to generalization phases. John’s peers had an average of 3.6 initiations (range 3–4) during baseline, 2.9 (range 1–5) during intervention, 0.8 (range 1–1) during maintenance, and 0.6 (range 0–1) during generalization. John’s peers slightly decreased responses during intervention, and then resumed baseline levels during maintenance and generalization. John’s peers had an average of 27.0 responses (range 20–40) during baseline, 19.8 responses (range 11–28) during intervention, 25.7 responses (range 17–33) during maintenance, and 31.3 responses (range 24–43) during generalization. During baseline, John performed far below his peer group for both initiations and responses. He moved closer to his peer group during intervention, though he still performed below the range of peer responses for initiations. During maintenance and generalization, he showed compa-
rable levels of initiations and responses to his peers.

**Brad and his peers.** Brad’s peers decreased initiations sharply from baseline to intervention phases and then gradually from intervention to generalization. Brad’s peers had an average of 4.5 initiations (range 2–11) during baseline, 1.8 (range 1–3) during intervention, 0.5 (range 0–1) during maintenance, and 0.6 (range 0–1) during generalization. Brad’s peers slightly decreased responses during intervention, and then increased responses above baseline levels during maintenance and generalization. Brad’s peers had an average of 4.5 initiations (range 2–11) during baseline, 1.8 (range 1–3) during intervention, 0.5 (range 0–1) during maintenance, and 0.6 (range 0–1) during generalization. Brad’s peers had an average of 19.4 responses (range 8–35) during baseline, 17.7 responses (range 9–30) during intervention, 28.0 responses (range 14–37) during maintenance, and 38.5 responses (range 19–42) during generalization. During baseline, Brad performed far below his peer group for initiations and responses. During intervention, maintenance and, generalization, he showed comparable levels of initiations and responses to his peers.

**Alton and his peers.** Alton’s peers decreased initiations sharply from baseline to intervention phases and then gradually from intervention to generalization. Alton’s peers had an average of 2.4 initiations (range 1–6) during baseline, 1.1 (range 0–2) during intervention, 0.3 (range 0–0) during maintenance, and 0.5 (range 0–1) during generalization. Alton’s peers demonstrated a gradual increase in responses across all phases of the study. Alton’s peers had an average of 16.6 responses (range 8–26) during baseline, 24.0 responses (range 10–46) during intervention, 26.0 responses (range 20–35) during maintenance, and 28.5 responses (range 19–40) during generalization. While Alton began baseline below his peer group, he showed comparable levels of initiations and responses to his peers by the end of baseline and throughout intervention, maintenance, and generalization.

**Evan and his peers.** Evan’s peers showed a slight decrease in initiations from baseline to generalization. Evan’s peers had an average of 1.2 initiations (range 0–2) during baseline, 0.3 (range 0–1) during intervention, 0.6 (range 0–1) during maintenance, and 0.2 (range 0–0) during generalization. Evan’s peers increased responses from baseline to intervention and then maintaining these levels during maintenance and generalization phases. Evan’s peers demonstrated an average of 20.3 responses (range 2–40) during baseline, 38.4 responses (range 15–62) during intervention, 32.7 responses (range 20–46) during maintenance, and 40.7 responses (range 29–58) during generalization. While Evan began baseline below his peer group, he showed comparable levels of initiations and responses to his peers by the end of baseline and throughout intervention, maintenance, and generalization.

**Social Validity**

For the teacher interviews, the average score was 4.35 out of 5 (range 3.75 to 5). Teachers reported that the intervention was a fair way to handle social skills difficulties, that they would suggest the intervention to other teachers, and that the need to learn social skills warrants use of the intervention. Peers in all groups reported that they liked the intervention and thought it was effective in increasing participants’ social skills. Overall, 12 of the 13 peers (92%) said that the intervention made the both the participants and themselves “better talkers.”

**Discussion**

The purpose of the study was to examine the extent to which peer mediation increased, maintained, and generalized social initiations and responses of students with ASD as compared to levels demonstrated by typically-developing peers. Probes to assess the persistence of generalization effects were collected and analyzed.

A comparison of Figures 2 and 3 indicates an inverse relationship between initiations and responses. As initiations decreased, responses increased throughout the study for participants and peer groups. By the end of the study, entire 10-minute sessions were sustained by one or two initiations. These results suggest that small numbers of initiations may result in lengthy discussions with numerous responses, and interventions that target responses may increase conversational turn-taking. According to DiSalvo and Oswald (2002), “many studies reported increases in responding or initiation by target children, but not both” (p. 204). Although the literature suggests that students with ASD particularly struggle with initia-
tions, interventions that target responses help to increase conversational-turn taking needed to sustain social interactions (Harjusola-Webb, Hubbell, & Bedesem, 2012).

A notable finding is that peers made fewer initiations after participating in peer training. Because intervention training emphasized both initiations and responses, it is possible that peers began to wait for participants to respond or became more sensitive to participants’ response attempts. Furthermore, the decrease in peer initiations brought levels of peer data in line with participant data. Researchers may want to compare average levels of initiations and responses in this study to studies that contain only typically-developing participants.

Two factors suggest that the intervention effectively increases responses. First, all participants increased response levels from baseline to intervention, with three of four participants (John, Brad, and Evan) having no overlapping data points from baseline to intervention. This demonstrates a functional relation between the intervention and the dependent variables. Second, participants increased responses to levels comparable to their peer groups, which some researchers consider necessary for a successful intervention (Ennis et al., 2013; Van Houten, 1979). Brad, Alton, and Evan demonstrated levels of responses equivalent to the levels of their peer groups during intervention, maintenance, and generalization phases. John performed slightly below his peer group during intervention but rose to the range of peer averages during maintenance and generalization.

**Generalization**

Koegel and Rincover (1977) identified three potential outcomes of generalization (generalization occurs, occurs and is maintained, or occurs but is not maintained). Visual analysis of Figure 1 shows that generalization of responses initially occurred for all participants, occurred and was maintained for two participants (John and Brad), and occurred but was not maintained by the other two participants (Alton and Evan). While still above baseline and intervention levels, data for Alton and Evan show a downward trend in responses during generalization. Evan’s data show stable levels of maintenance, indicating that for Evan, the intervention maintained but generalization across materials was not maintained. While still above baseline and intervention levels, Alton’s maintenance and generalization data decreased in trend throughout maintenance and generalization phases. This indicates that for Alton, the intervention may not continue to be maintained, and generalization across materials also may not continue to be maintained.

Analyzing maintenance and generalization as interrelated gives researchers a new lens through which to evaluate the success of an intervention. While results from this study support previous research that students maintain and generalize social gains through peer mediation (see Chan et al., 2009; Watkins et al., 2015), literature reviews typically treat generalization as a binary variable (i.e., yes/no generalization was assessed or yes/no generalization occurred). In this study, analysis indicating that generalization wasn’t maintained for two participants enabled researchers to predict future performance and customize intervention plans.

**Social Validity**

As a measure of social validity, each participant’s levels of social interaction were compared to the peer norms of his group (Ennis et al., 2013). In addition to using peer norms, social skills assessments and treatment acceptability questionnaires were used to comprehensively determine the social validity of the intervention (Ennis et al., 2013; Wolf, 1978). Cumulatively, these results suggest that the teachers and peers in this study considered peer mediation to be a socially valid intervention. Future researchers may also want to interview participants about treatment acceptability as a way for participants to self-evaluate and reflect upon their progress.

**Implications for Practice**

When conducting peer-mediated social skills interventions, researchers and practitioners should keep several criteria in mind. First, interventions should be conducted in natural settings with typically-developing peers when possi-
able (Harrower & Dunlap, 2001). Second, research indicates that careful selection of target skills (Gresham, Elliott, & Kettler, 2010), peers (Van Houten, 1979), materials (Martin, Brady, & Williams, 1991), and activities (Odom & Strain, 1984) may augment the strength of peer-mediated interventions. Third, peer norms may be used to evaluate student progress and the effectiveness of the intervention (Van Houten, 1979). Finally, the persistence of generalization effects may be analyzed to develop a comprehensive intervention plan (Koegel & Rincover, 1977).

Limitations and Future Research

Several issues affect the interpretation of results in this study. The authors defined the dependent variables in the study to include only verbal initiations and responses, which may give an incomplete picture of overall communication attempts. Most initiations have numerous acceptable responses that can be either verbal or non-verbal. For example, when a student asks a peer to hand him the spinner, the response could be verbal (“Sure” or “Here it is” or “No, it’s my turn”) or nonverbal (nodding head or simply handing over the spinner). To respect the natural classroom environment and encourage authentic peer conversation, participants and peers were allowed to respond naturally, which often included both verbal and non-verbal responses, and move freely in their small groups. Sometimes students had their backs to the camera, and nonverbal communication attempts could not be reliably captured.

Researchers used peer selection criteria commonly used in social skills studies (Watkins et al., 2015). Although participants increased responses to levels within the range of their peer groups, it is unclear if the peer groups performed above, at, or below the levels of social skills typically demonstrated by students of the same age (Van Houten, 1979). Future researchers may want to screen potential peers by administering the SSIS (Gresham & Elliott, 2008) to evaluate how peers perform as compared to a larger sample (Greenwood et al., 1981; Gresham et al., 2010).

Materials selection is a highly variable factor in social skills interventions that deserves further investigation. In this study, the level and trend of participant data often corresponded with the level and trend of their peer group data (see Figure 2). Materials selection may provide one explanation for the similarities in performance between participants and their peer groups. Educational games in this study required basic game-play structure (take turns, follow game rules, or contribute pieces to a common board). However, levels of social interaction may have been influenced by unknown factors, such as students’ prior exposure to specific games, games with similar structure, or academic skills used in the games. Use of familiar or highly-preferred games may have contributed to high levels of social interactions for participants and their peer groups in some sessions. Future researchers may want to explore how different materials affect the performance of social skills for specific groups of students (Morrison et al., 2001).

Conclusion

The purpose of this study was to examine the extent to which peer mediation increased, maintained, and generalized social initiations and responses of students with ASD across academic settings as compared to their typically-developing peers. Results indicate that peer mediation increased, maintained, and generalized responses for all participants but did not significantly impact initiations. Additionally, all participants increased responses to levels within the range of responses demonstrated by their peer groups. These outcomes suggest that peer mediation is an effective intervention to increase, maintain and generalize social skills and is a socially valid intervention for participants in the study.

References


McConnell, S. R. (2002). Interventions to facilitate social interaction for young children with autism: Review of available research and recommendations for educational intervention and future re-


Received: 6 April 2017
Initial Acceptance: 4 June 2017
Final Acceptance: 15 August 2017