Abstract: This meta-analysis investigated the efficacy of peer-mediated interventions for promoting social interactions among children from birth to eight years of age diagnosed with ASD. Forty-five single-subject design studies were analyzed and the effect sizes were calculated by the regression model developed by Allison and Gorman (1993). The overall effect sizes suggest that peer-mediated interventions were highly effective. Further categorical comparisons suggest that these interventions were more effective in enhancing social responses in younger boys, when older male siblings served as interventionists, when the interventions took place in the home, when peer modeling was used, and when consideration was given to maintenance and generalization across participants, behaviors and activities, and in involving collaboration among all researchers, peers/siblings, school staff, and parents/families.

Evidence-based practice in the education and treatment of children and youth with autism enables special education professionals to validate their practices in accordance with the existing body of scientific evidence found in the literature (Wheeler, 2007). Yet this is a challenge for the field given the heterogeneity of the participants and the varying educational contexts that serve children diagnosed with Autism Spectrum Disorders (ASD) (Odom et al., 2005; Wheeler). Given that the prevalence of ASD is increasing at a dramatic rate, and that social interaction is a core deficit associated with ASD, the urgency for validating evidence-based practice in the study of social competence among young children with autism is critically important. The field of special education has embraced the concept of evidence-based practice yet researchers and practitioners have failed to operationalize this construct within the practice of research and delivery of educational services and supports. Thus, it is vital for professionals to systematically synthesize extant research within an evidence-based framework, so that basic and applied research can be translated to applied practice to better assist practitioners in the design and delivery of efficacious interventions and supports to children with ASD and their families (Wheeler).

Evidence-based Practices

The gap between research to practice in the field of education has been a source of debate for many decades as policy makers have argued that many practices lack the empirical efficacy needed to substantiate their claim. The most recent legislation, the No Child Left Behind Act (NCLB), draws attention to the need for greater levels of accountability thus requiring teachers to use scientifically proven practices in their classrooms (Odom et al., 2005; U.S. Department of Education, 2003). In spite of this recent mandate, there continues to be a disparity within the field of special education as to the fidelity of how these practices are implemented within classrooms among learners with disabilities. This is perhaps due in part to the complexity of this field, including the variability of the participants and the varying degrees of educational

Correspondence concerning this article should be addressed to Jie Zhang, Education and Human Development, The College at Brockport, State University of New York, 350 New Campus Drive, Brockport, NY 14420. E-mail: jzhang@brockport.edu
contexts that serve children with disabilities (Odom et al.; Wheeler, 2007). Odom and colleagues assert that the field of special education needs to develop specific guidelines for specifying types and levels of evidence needed to identify a practice as evidence-based and effective. These have yet to be fully operationalized in the literature. Until professionals have agreed on standards for determining evidence-based practices there will continue to be problems caused by misinformed practice. Researchers can assist in minimizing these irregularities through the systematic synthesis of extant research to ascertain the existing knowledge base and the gaps that exist between research and practice (Odom et al.; Wheeler). Furthermore, the knowledge from the current literature can be used to inform practice and scientifically proven evidence can be adopted as the appropriate basis for selecting these practices (Odom et al.).

Meta-Analysis for Single-Subject Studies

Single-subject research is a rigorous, scientific methodology used to define basic principles of behavior and to establish evidence-based practices by documenting functional relationships between independent and dependent variables (Horner et al., 2005). It provides systematic and detailed analysis of individuals and has proven especially relevant for defining educational practices at the individual learner level (Horner et al.). Thus, it plays an important role in identifying evidence-based practice in the field of special education. However, the criticism often leveled against single subject methodology is that the sample sizes used in these studies are often too small to be generalized to a larger population (Wellen, 1998). Synthesis of single-subject studies is one way to overcome this generalization problem since the procedures help determine whether a specific intervention is consistently effective in the change of the target behavior(s). In addition, if employed appropriately, the information from the synthesis of single-subject studies can contribute potentially to the field of special education (Scruggs, Mastropieri, & Casto, 1987).

Meta-analyses for single-subject studies can be used to develop guidelines for evidence-based practice and to design new studies because of the following reasons: (a) The aggregation of findings from a large number of single-subject studies can get a large enough sample to strengthen the conclusions about the practical implications to practitioners (Gingerich, 1984). (b) Information is taken from graphs and an unbiased synthesis of the empirical data can produce a more accurate estimation of the impact of the intervention (Davies & Crombie, 2001; Wellen, 1998). (c) The compilation of findings can identify factors contributing to the effectiveness, so that an intervention can be tailored more specifically to the unique characteristics of the participants and situation (Gingerich; Wheeler, 2007). (d) The method of coding can point up the failure to report some important variables in some studies (White, Rusch, Kazdin, & Hartmann, 1989) and also identify gaps in the existing research literature (Wheeler).

Autism Spectrum Disorders

Perhaps the greatest area of need for conducting such critical analyses of existing research lies in the area of ASD due to its dramatically increasing prevalence. Once believed to be a low-incidence disorder, ASD was estimated to occur in 4–5 children out of 10,000 prior to 1985 (Byrd et al., 2002). However, it is far more common than previously thought. It was estimated that the prevalence of ASD increased approximately four times in one decade, from 1 in 1,333 children (7.5 per 10,000 children) among children born in the mid-1980s to 1 in 323 children (31.2 per 10,000 children) among children born in the late-1990s (California Health and Human Services Agency, 2003). In 2004, the Department of Health and Human Services and the American Academy of Pediatrics issued an “Autism ALARM” claiming that 1 in 166 children (60 per 10,000 children) had ASD. Nowadays, the prevalence rates of children with ASD have reached 1 in 150 children (66 per 10,000 children) (Centers for Disease Control and Prevention, 2007). Accordingly, the number of children receiving services for ASD is reportedly on the rise. ASD is now recognized to be more prevalent in childhood than diabetes, cancer, spina bifida, and Down syndrome (Filipek et al., 1999).

Given that the prevalence of ASD is increas-
ing at a dramatic rate, the urgency for validating evidence-based practice is critically important. One area of vital importance is social competence. Since social interaction is a core deficit associated with ASD, requests on the effectiveness of the interventions aimed at increasing social interactions among young children with ASD have been quite striking (Goldstein, 2002). There is, however, a question on whether professionals have enough data to validate their practices as effectiveness (Wheeler, 2007). Furthermore, challenges exist in the implementation of evidence-based practice concerning the service for children with ASD because of: (a) the irregularities in the ability of special education to adopt the evidence-based practice, (b) a broad range of practices across educational settings exists in the delivery of services to these children including those which are not evidence-based, untested, ineffective, and sometimes even harmful practices (Bellini & Akullian, 2007; Wheeler), (c) the heterogeneity of the participants and the educational contexts that serve these children, and (d) limited sample size resulting in an inability to generalize findings (Wheeler). As a result, there is an urgent need for meta-analyses of the existing literature on the effectiveness of various treatments for children with ASD, in this case interventions directed towards increasing social competence in young children with ASD.

**Peer-Mediated Interventions for Children with Autism**

Peer-mediated interventions provide typically developing peers with such activities as social skills training, prompts and praise in social play situations which are designed to model, reinforce, and promote appropriate social interactions and social skills for children with disabilities (McConnell, 2002). In peer-mediated interventions, adults may facilitate and monitor the intervention from close by although they never intervene directly with the target children (Odom & Strain, 1984). Unlike adult-mediated approaches, peer-mediated interventions facilitate learning in natural social contexts with peers or siblings by precluding the additional steps required to transfer learning from adults to peers in natural social contexts (McConnell).

Utley, Mortweet, and Greenwood (1997) classify peer-mediated interventions into six categories. They state that peer-mediated interventions consist of (a) peer modeling, (b) peer initiation training, (c) peer monitoring, (d) peer networking, (e) peer tutoring, and (f) group-oriented contingencies. Peer modeling includes peer-proximity and peer-pairing, which rely on the inclusion of a socially competent peer to demonstrate appropriate behavior for a child with ASD to imitate. Peer modeling also includes video modeling, which shows a video with a socially competent peer teaching appropriate behaviors to a child with ASD (Utley et al.). Peer initiation training typically requires a teacher to train peers how to evoke and maintain desired social behaviors from a child with ASD by establishing eye contact, suggesting play activities, initiating conversation, offering or asking for help, describing ongoing social interactions, expanding the content of the target child’s speech, or demonstrating affection (Utley et al.). Peer monitoring uses a buddy system or role-playing to minimize teachers’ disciplinary and supervisory responsibilities. With peer monitoring, children with ASD are taught to function independently from teacher’s monitoring and management (Utley et al.). Peer networking aims to create a support system of friends for children with ASD in a natural social context. Socially competent peers are instructed to prompt and encourage social responses from target children, in addition to model and reinforce appropriate social behaviors (Utley et al.). Peer tutoring uses peers as one-on-one teachers to provide individualized instruction, practice, repetition, and concept clarification. It is used to increase opportunities to respond, academic engagement, and relevant academic behaviors for specific academic tasks (Utley et al.). With group-oriented contingencies, peers are trained as social change agents in a natural educational environment. When working for a common goal or reward, children in groups provide one another with natural social prompts and consequences. They distribute points or reinforcers, record data, give instructions, impose contingencies, provide assistance through cooperative reinforcement contingencies, and earn reinforcers depending to some extent on their behavior (Utley et al.).
DiSalvo and Oswald (2002) point out that integrated play groups, peer buddy and peer tutoring, and group-oriented contingencies arrange the situation or contingencies, enhance the likelihood of attention to peer models from children with autism and thereby, promote peer interactions. Peer networks, pivotal response training, and peer initiation training teach peers specific social skills, and make it easier and more rewarding for peers to interact with children with autism (DiSalvo & Oswald). McConnell (2002) reviewed 55 experimental research articles on interventions to facilitate social interaction for children with autism younger than nine years of age through 2000. Among the 55 studies, about thirty peer-mediated interventions were reviewed: (a) interventions to increase peer social initiations, (b) interventions to increase social and communicative interactions, and (c) peer-mediated incidental teaching and more structured peer tutoring (McConnell). Hwang and Hughes (2000) assessed 16 empirical studies from 1981 through 1997 to investigate the effects of social interactive interventions on early social communicative skills of young children with autism, aged two to twelve, by increasing their role as initiator of social interactions. Two studies used peer-mediated intervention. Results indicated prolonged social interactions (McGee, Almeida, Sulzer-Azaroff, & Feldman, 1992; Pierce & Schreibman, 1995) and increased engagement in joint attention (Hwang & Hughes; Pierce & Schreibman). Odom and his colleagues (2003) examined 37 single-subject design studies from 1990 to 2002 to investigate the scientific evidence of intervention effectiveness for young children with autism. As the result of this review, peer-mediated interventions were categorized as emerging and effective practices (Odom et al.).

The purpose of this study was to determine whether peer-mediated interventions were effective in improving the social interactions among young children with ASD. In addition, the study was to provide an analysis of the effects of the interventions by treatment variables: (a) target children’s characteristics; (b) interventionists’ characteristics; and (c) features of interventions. This study provides a better understanding concerning the utility of peer-mediated interventions among children under eight years of age, who were diagnosed with autism, for promoting social interactions. This study also refines evidence-based peer-mediated intervention practices for these children by providing detailed integrated findings through meta-analysis of individual single-subject studies.

Method

Criteria for Inclusion

Studies selected met six inclusion criteria as follows:

1. The study used peer-mediated interventions;
2. One of the purposes of the study was to enhance social interactions for children with ASD;
3. The target children were under eight years of age diagnosed with ASD;
4. The study employed single-subject designs that provided at least three data points for pre-intervention and three data points for post-intervention phases via detailed graphs;
5. Dependent variables included observed social interactions between young children with ASD and their peers; and
6. All the articles were published in peer-reviewed English language journals between 1977 and 2006.

Search Procedure

Searches were carried out to obtain articles for inclusion in this meta-analysis. First, an electronic search was conducted among all peer-reviewed English language journals published from 1977 to 2006 using ERIC and PsycINFO search database. With one keyword from each of the three categories, combinations of three keyword entries were used to select studies for the present meta-analysis: (a) autism, autistic; (b) social development, social interaction, social behavior, social competence, social skills, peer relationship, peer relation, socialization, friendship, friend; and (c) children, young children, early intervention. The total number of combined keyword sets was 60 (2x10x3). The search process resulted 2,670 articles. After eliminating 2,045 duplicates and exclud-
ing all irrelevant articles (e.g., reviews and position papers), a total number of 94 studies were retained.

Second, a hand search started with journal articles published in 1977 or with the first issue if the journal was founded more recently. The journals included Journal of Applied Behavior Analysis, Focus on Autism and Other Developmental Disabilities, and Topics in Early Childhood Special Education. Additionally, relevant studies found in the reference section of the reviewed articles were located and chosen according to the selection criteria. The second search resulted in 45 articles, 118 children, 190 graphs of observations, and 6152 observation data points from 19 journals.

Data Coding
A coding sheet was developed and used during both the data coding and double-coding procedures by the two authors. Relevant data from each selected study were coded using a systematic set of rules and procedures across the following categories: (a) target children’s characteristics, including age, gender, and diagnosis; (b) interventionists’ characteristics, including relative age, gender, and relation with the target children; and (c) features of the interventions, including settings, intervention types, target behaviors, with the consideration of maintenance or not, with the consideration of generalization or not, length of intervention, intensity of intervention, and involvement of researchers, peers/siblings, school staff and parents/families. Length of intervention was calculated by the length of intervention per session multiplied by the total number of sessions, while intensity of intervention was calculated by the length of intervention per session multiplied by the number of sessions per week. If a study reported the actual length of the intervention, the data was recorded. If, however, the length of the intervention was not reported in a study, it was calculated by using the length of the observational sample from the graphs. Data were transferred from coding sheets to Excel and then SPSS data sheets.

Inter-Rater Reliability
Inter-rater reliability for the coding procedure was determined through double-coding by the two authors of this meta-analysis. Twenty studies (44%) from the selected articles were randomly chosen and independently double-coded by the secondary author, Professor of Special Education and the primary author’s doctoral advisor. Additionally, given the complex nature of Allison and Gorman’s (1993) regression model, inter-rater reliability for the effect sizes analysis was determined through independent calculation by the primary author’s colleague, a senior graduate student of Special Education. The regression model was initially reviewed and then the effect sizes were computed independently for all 45 studies. Inter-rater reliability was calculated by dividing the number of agreements by the total number of agreements and disagreements and multiplied by 100%. The mean inter-rater reliabilities were 97.6% for the coding procedure and 100% for the effect size calculation.

Analyses of Data
Descriptive data of the selected studies were analyzed by calculating both the frequency and percentage for each of the variables from the coding sheet. The percentages were calculated by dividing the number of items in a subset by the total number of items in that variable. The regression model refined by Allison and Gorman (1993) was used to calculate the effect sizes. That model was chosen because it accounts for the natural trend of change by time in observations, i.e., it takes into account improvement in the baseline phase to avoid overestimating the effects of treatment (Allison & Gorman). To produce an effect size that excludes any improvement that may occur without intervention, it uses the trend of baseline observations while adjusts the treatment observations by subtracting the predicted values from the actual values in the treatment. It also considers the need to represent negative effects with a negative value of the correlation. In addition, it addresses the problem posed by changes in level and slope that occur in opposite directions (Allison & Gorman; Wellen, 1998). Thus, if the trend toward improvement in the baseline phase increases in the treatment phase, or the upward sloping line in the baseline phase gets steeper in the treatment phase, the method
considers that change and calculates a greater effect size, and vice versa (Wellen).

Each usable graph of every selected child in the included studies was analyzed:

1. AB designs: the observations in the baseline A were compared to observations in the intervention phase B;
2. A1B1A2 designs: only the observations in the first baseline A1 were compared to observations in the intervention phase B1;
3. A1B1A2B2 designs: only the observations in the first baseline A1 were compared to observations in the last intervention phase B2.

As Allison and Gorman (1993) state, typically, researchers terminate their first treatment B1 before achieving maximum effects with the concern that the behavior will not return to the baseline A1 if treatment is carried “too far.” Thus, the maximum effects attempt to be achieved in the last intervention phase B2. In addition, the second baseline A2 always returns to the first baseline level of performance. Accordingly, the comparison of A1 and B2 is thought to yield the most valid estimate of treatment effects (Allison & Gorman);

4. Multiple baseline designs across behaviors: only the graph with the social interactions with peers as the target behavior was used;
5. Multiple baseline designs across settings: all data points from all of the graphs were used;
6. Multiple baseline designs across participants: every young child with autism was regarded as a separate AB design;
7. Multiple probe designs: similar to the multiple baseline designs;
8. Baseline followed by alternating treatments designs, baseline followed by alternating treatments and a final treatment phase designs, and adapted alternating treatments designs: the observations in the baseline A were compared to observations in the peer-mediated intervention phase;
9. Simultaneous treatment designs: the observations in the baseline A were compared to observations in the peer-mediated intervention phase;
10. Changing criterion designs: only the observations in the first and only baseline A were compared to observations in the last intervention phase.

After collecting the data from each selected graph, the effect sizes were calculated. First, among all the selected graphs, the intervention effect sizes (IES) by the measure of frequency were calculated between the baseline phase and the intervention phase for each selected child in every included study. Then, the overall intervention effect size (OIES) for frequency was averaged based on the IESs of the individual target children. Similarly, the follow-up effect sizes (FES) and generalization effect sizes (GES) by the measure of frequency were calculated between the baseline phase and the follow-up phase, or between the baseline phase and generalization phase, respectively, for each selected child. The overall follow-up effect size (OFES) and overall generalization effect size (OGES) for frequency was averaged based on FESs and GESs of the individual target children in all selected studies. Likewise, OIES, OFES, and OGES by the measure of duration were averaged based on IESs, FESs, and GESs for duration of the individual target children. In addition, the intervention effect sizes (IES) for frequency due to each variable from the coding sheet were used to calculate for each child in every selected study for further categorical comparisons.

The effect sizes were weighted by the number of observation data points in each graph (Wellen, 1998). When the predicted values went beyond the possible limits, they were set back at the natural limits (Campbell, 2003). For example, when the predicted values of frequency were negative, they were set at 0. Similarly, when the predicted values of frequency were above 100, they were set at 100. If two or more sets of data were collected for one child in a selected study, only the set with the largest number of data points was used to guarantee one effect size for each selected child and also to avoid the average of the effect sizes for the child. If one target child had more than one pair of largest number of
observation data points in the baseline and intervention phases, the pair was randomly chosen. For example, if the intervention used multiple baseline design across settings and was conducted for one child across three different settings with the same number of data points in the three baselines and the same number of data points in the three intervention phases, then the pair was randomly chosen. Furthermore, means and standard deviations were calculated, and the effect sizes over three standard deviations were eliminated as outliers.

The one-sample \( t \) test was used to determine whether or not the overall effect sizes were significantly different from zero. In addition, a one-way ANOVA was used to see whether there was any significant difference within different variables. Once statistical significance was found, multiple comparisons were conducted to determine whether there were significant contrasts. Effect sizes of this meta-analysis were defined according to Cohen’s \( d \) (1988) standard: (a) \( ES = .2 \) is considered as a small effect; (b) \( ES = .5 \) a medium effect; and (c) \( ES = .8 \) a large effect.

## Results

### Articles Selected for Inclusion

The selected studies were published from 1978 to 2006. Among them, twenty studies (40%) were published between 1992 and 1997. Furthermore, the 45 studies selected in this meta-analysis were published in 19 journals. Twelve studies (26.7%) were published in the *Journal of Applied Behavior Analysis*, five studies (11.1%) in the *Journal of Autism and Developmental Disabilities*, and four studies (8.9%) in the *Journal of the Association for Persons with Severe Handicaps*. Table 1 presents the frequency and percentage of the selected studies published by journal.

### Overall Effect Sizes

Table 2 presents the overall effect sizes. Results suggest that peer-mediated interventions for promoting social interactions among young children with ASD were highly effective; and this effectiveness lasted across time, different settings, participants, target behaviors, or activities (\( p < .01 \)).
Analyses by Participants’ Characteristics

The total number of studies analyzed by the target children’s age was larger than 45 because some studies included children at different age level. Twenty-six studies included target children from 72 to 97 months (42.6%), while one study included a target child under three years of age (1.6%). Forty-four studies included boys as the target children (78.6%), compared to 12 studies that included girls (21.4%). Forty-four studies included target children diagnosed with autism (88%). No study included target children diagnosed with Rett Syndrome or Childhood Disintegrative Disorder (CDD). Interventionists’ age was coded relatively compared to the target children’s age. Twenty-one studies included interventionists older (32.8%). Eighteen studies did not specify the gender of the interventionists (35.3%). The majority of the selected studies (n = 39) included peers as interventionists (84.8%).

The IES by age 0–35 months was not calculated due to the constant variables in the baseline phase. Since the number of data of the IES by autism was much larger than those of IES by Asperger’s or by PDD-NOS, these three IESs were not compared. Likewise, the IES by both siblings and peers as the interventionists was not compared due to the large difference between the number of data.

Results of the effect sizes suggest that peer-mediated interventions for promoting social interactions among young children with ASD were highly effective across target children’s: (a) age, (b) gender, and (c) diagnosis. They were also highly effective: (d) when having interventionists older, at the same age, or with the combined age levels; (e) across different genders of the interventionists; and (f) across different relation between the interventionists and the target children. The intervention was more effective (a) in younger children (p < .01), (b) in boys (p < .01); or if the interventionists were (c) older or at different age levels (p < .01), (d) boys (p < .01), and (e) siblings (p < .01). Table 3 provides data related to the participants’ characteristics, such as the target children’s age, gender, diagnosis, in addition to the interventionists’ relative age, gender, and their relation with the target children.

Analyses by Features of the Interventions

Twenty-two studies were carried out in integrated classrooms (44.9%), while one study was conducted in a clinical setting (2%). Thirteen studies used more than one intervention (27.2%). More than half of the studies (n = 37) studied social interactions combined with both initiation and response (62.7%). Eighteen studies reported maintenance (38.3%), sixteen studies reported generalization across settings (34.8%), and all 45 studies (100%) carried out generalization across participants: some interventions were implemented to different target children, and others with different interventionists. Twenty-four studies reported generalization across behaviors (53.3%), twenty-one studies reported generalization across activities (46.7%), and more than a half of the studies (n = 30) involved researchers, peers/siblings, and school staff in the intervention (63.8%). The length of intervention was calculated by the length of intervention per session multiplied by the total number of sessions. Twenty-one studies performed the intervention for less than one hour (22.6%) and another 21 studies between one and two hours (22.6%). The intensity of intervention was calculated by the length of intervention per session multiplied by the number of sessions per week. Thirteen studies performed the intervention less than half an hour a week (28.9%), while twelve studies did not specify the intensity of the intervention (26.7%).

The IES by clinics was negatively high. It could not, however, be generalized since only one study occurred in a clinical setting. There was no statistical significant difference between the participants’ characteristics, such as the target children’s age, gender, diagnosis, in addition to the interventionists’ relative age, gender, and their relation with the target children.

<table>
<thead>
<tr>
<th>Overall Effect Sizes</th>
<th>N</th>
<th>ES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall IES for Frequency</td>
<td>1880</td>
<td>1.46**</td>
</tr>
<tr>
<td>Overall FES for Frequency</td>
<td>446</td>
<td>1.49**</td>
</tr>
<tr>
<td>Overall GES for Frequency</td>
<td>219</td>
<td>1.51**</td>
</tr>
<tr>
<td>Overall IES for Duration</td>
<td>401</td>
<td>1.27**</td>
</tr>
<tr>
<td>Overall FES for Duration</td>
<td>79</td>
<td>2.44**</td>
</tr>
</tbody>
</table>

Note: ** indicates ES is significantly different from 0 (p < .01).
The IES by the involvement of researchers, peers/siblings, and parents/families was not compared, due to the limited number of studies and limited number of observation data. All ten IESs by the intervention length were compared, and the IES by 7–8 hours duration was significant different from other IES \((p < .01)\). When comparing continuously, however, there was no bivariate correlation between the IES and length of intervention \((p > .05)\). Similarly, all six IESs by the intervention intensity were compared, and there were significantly differences between some of them \((p < .05)\). There was, however, no bivariate correlation between the IES and continuous intensity duration \((p > .05)\).

Results of the effect sizes by the characteristics of the interventions from this study suggest that peer-mediated interventions among young children with ASD were highly effective: (a) across the settings, except clinics; (b) across different intervention types except for peer networking; (c) for promoting social responses and social interactions; (d) whether or not with the consideration of maintenance; (e) whether or not with the consideration of generalization; (f) across different involvement of participants; (g) across different length of intervention except for 4–5 hours duration; and (h) across all the intensity of the intervention. Furthermore, the intervention was more effective if the intervention: (a) took place at home \((p < .01)\); (b) used peer modeling \((p < .01)\); (c) aimed to enhance social response \((p < .01)\); (d) considered maintenance \((p < .05)\); (e) considered generalization across settings \((p < .05)\).
alization across participants, across behaviors, or across activities ($p < .01$); and (f) involved all researchers, peers/siblings, school staff, and parents/families ($p < .01$). Table 4 provides data related to the settings, intervention types, and target behaviors, maintenance, generalization, collaboration, the length and intensity of intervention.

**Conclusions**

A meta-analysis was conducted in 45 single-subject design studies from 19 journals between 1977 and 2006. The overall effect sizes suggest that peer-mediated interventions were highly effective among children under eight years of age diagnosed with ASD for promoting social interactions. Further categorical comparisons suggest that these interventions were more effective in enhancing social responses in younger boys, when older male siblings served as interventionists, when the interventions took place in the home, when peer modeling was used, and when consideration was given to maintenance and generalization across participants, behaviors and activities, and in involving collaboration among all researchers, peers/siblings, school staff, and parents/families.

**Limitations**

One limitation of the presented meta-analysis stems from the inclusion/exclusion criteria used to select the studies. Only the studies published in peer-reviewed journals were included. Accordingly, this meta-analysis was biased in favor of published research compared to unpublished studies. Horner, Carr, Strain, Todd, and Reed (2002) state that due to a potential bias imposed by publication procedures, studies with positive results are more likely to be published than studies without or with negative effects. Such practices lead to a possible inflated estimate of effect size represented by published studies (Horner et al.). Therefore, the submission and publication of studies with negative or ineffective findings should be encouraged.

Allison and Gorman’s (1993) regression model was used to perform this meta-analysis. The model accounts for the natural trend of change in baseline to adjust the treatment observations, so it may rely too heavily on the estimation of trends from the baseline phase. Thus, the method may lead to inaccurate calculations of effect sizes when the number of data points in baseline is limited (Wellen, 1998).

Cohen’s $d$ standards were used to define the effect sizes of this meta-analysis. However, instead of within subject comparisons in single-subject design studies, they were meant for between group comparisons in group designs. In many meta-analyses using single-subject design studies, effect sizes usually occurred larger than those in group design studies. Thus, the interpretation of the results of this study may change if the standards of the effect sizes change.

Another limitation derives from the use of a ruler to measure data points from the selected graphs during the process of data collection. The value of each data point in some selected studies had to be estimated because the graphs were too small to read exactly. To make the process of meta-analyses more efficient and to obtain more accurate information, single-subject studies in the future should provide more complete information: the original data in the form of graphs and in numbers, the number of baseline and treatment observation data points, and a more detailed description of the intervention variables.

Even though the effect sizes suggested that the intervention was more effective if the interventionists were siblings of the target children and if the intervention took place at home, consideration should be given regarding the confound between the siblings as interventionists and home-based intervention, since most home-based interventions included siblings’ participation. For example, among the five studies which took place at home, only one child, Huang, in one study by Yang, Wolfsberg, Wu, and Hwu (2003) had both peer and siblings as the interventionists. Similarly, among the five studies which siblings served as the interventionists, only one study by Baker (2000) took place in playrooms at the university instead of home. The other four studies took place at home while had sibling serving as the interventionists (Coe, Matson, Craigie, & Gossesn, 1991; Strain &
### TABLE 4
Data Related to the Features of the Interventions

<table>
<thead>
<tr>
<th>Variables</th>
<th>Frequency</th>
<th>Percent (%)</th>
<th>N</th>
<th>ES</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Intervention Setting</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Integrated Class</td>
<td>22</td>
<td>44.9</td>
<td>987</td>
<td>1.13</td>
</tr>
<tr>
<td>Segregated Class</td>
<td>6</td>
<td>12.2</td>
<td>192</td>
<td>1.86</td>
</tr>
<tr>
<td>Home</td>
<td>5</td>
<td>10.2</td>
<td>197</td>
<td>2.23**</td>
</tr>
<tr>
<td>Clinic</td>
<td>1</td>
<td>2.0</td>
<td>19</td>
<td>-1.56</td>
</tr>
<tr>
<td>Other</td>
<td>15</td>
<td>30.6</td>
<td>485</td>
<td>1.78</td>
</tr>
<tr>
<td><strong>Independent Variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peer Modeling</td>
<td>6</td>
<td>12.8</td>
<td>127</td>
<td>3.16**</td>
</tr>
<tr>
<td>Peer Initiation Training</td>
<td>10</td>
<td>21.3</td>
<td>322</td>
<td>0.97</td>
</tr>
<tr>
<td>Peer Monitoring</td>
<td>3</td>
<td>6.4</td>
<td>204</td>
<td>1.45</td>
</tr>
<tr>
<td>Peer Networking</td>
<td>3</td>
<td>6.4</td>
<td>133</td>
<td>0.62</td>
</tr>
<tr>
<td>Peer Tutoring</td>
<td>3</td>
<td>6.4</td>
<td>50</td>
<td>1.37</td>
</tr>
<tr>
<td>Group-Oriented Contingency</td>
<td>9</td>
<td>19.1</td>
<td>435</td>
<td>1.00</td>
</tr>
<tr>
<td>Combination</td>
<td>13</td>
<td>27.7</td>
<td>609</td>
<td>1.89</td>
</tr>
<tr>
<td><strong>Dependent Variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social Initiation</td>
<td>13</td>
<td>22.0</td>
<td>265</td>
<td>0.71</td>
</tr>
<tr>
<td>Social Response</td>
<td>9</td>
<td>15.3</td>
<td>220</td>
<td>3.23**</td>
</tr>
<tr>
<td>Combination of Social Interaction</td>
<td>37</td>
<td>62.7</td>
<td>1395</td>
<td>1.33</td>
</tr>
<tr>
<td><strong>Intervention Maintenance</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>18</td>
<td>38.3</td>
<td>789</td>
<td>1.54*</td>
</tr>
<tr>
<td>No</td>
<td>29</td>
<td>61.7</td>
<td>1091</td>
<td>1.41</td>
</tr>
<tr>
<td><strong>Intervention Generalization Across Settings</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>16</td>
<td>34.8</td>
<td>577</td>
<td>1.37</td>
</tr>
<tr>
<td>No</td>
<td>30</td>
<td>65.2</td>
<td>1303</td>
<td>1.50</td>
</tr>
<tr>
<td><strong>Intervention Generalization Across Participants</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>45</td>
<td>100.0</td>
<td>1880</td>
<td>1.46**</td>
</tr>
<tr>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Intervention Generalization Across Behaviors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>24</td>
<td>53.3</td>
<td>954</td>
<td>1.76**</td>
</tr>
<tr>
<td>No</td>
<td>21</td>
<td>46.7</td>
<td>926</td>
<td>1.16</td>
</tr>
<tr>
<td><strong>Intervention Generalization Across Activities</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>21</td>
<td>46.7</td>
<td>986</td>
<td>1.67**</td>
</tr>
<tr>
<td>No</td>
<td>24</td>
<td>53.3</td>
<td>894</td>
<td>1.25</td>
</tr>
<tr>
<td><strong>Involvement of Researchers, Peers/Siblings, School Staff and Parents/Families</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Researchers &amp; Peers/Siblings</td>
<td>7</td>
<td>14.9</td>
<td>186</td>
<td>1.98</td>
</tr>
<tr>
<td>Researchers &amp; Peers/Siblings, &amp; School Staff</td>
<td>30</td>
<td>63.8</td>
<td>1376</td>
<td>1.31</td>
</tr>
<tr>
<td>Researchers &amp; Peers/Siblings, &amp; Parents/Families</td>
<td>2</td>
<td>4.3</td>
<td>57</td>
<td>0.63</td>
</tr>
<tr>
<td>Researchers &amp; Peers/Siblings, School Staff, &amp; Parents/Families</td>
<td>8</td>
<td>17.0</td>
<td>261</td>
<td>2.06**</td>
</tr>
<tr>
<td><strong>Length of Intervention (minutes)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10–60</td>
<td>21</td>
<td>22.6</td>
<td>443</td>
<td>1.75</td>
</tr>
<tr>
<td>61–120</td>
<td>21</td>
<td>22.6</td>
<td>358</td>
<td>1.82</td>
</tr>
<tr>
<td>121–180</td>
<td>13</td>
<td>14.0</td>
<td>195</td>
<td>0.90</td>
</tr>
<tr>
<td>181–240</td>
<td>8</td>
<td>8.6</td>
<td>152</td>
<td>1.23</td>
</tr>
<tr>
<td>241–300</td>
<td>5</td>
<td>5.4</td>
<td>146</td>
<td>0.33</td>
</tr>
<tr>
<td>301–360</td>
<td>3</td>
<td>3.2</td>
<td>64</td>
<td>0.92</td>
</tr>
<tr>
<td>361–420</td>
<td>6</td>
<td>6.5</td>
<td>164</td>
<td>0.95</td>
</tr>
<tr>
<td>421–480</td>
<td>3</td>
<td>3.2</td>
<td>69</td>
<td>4.02**</td>
</tr>
<tr>
<td>481–540</td>
<td>5</td>
<td>5.4</td>
<td>143</td>
<td>1.46</td>
</tr>
<tr>
<td>&gt;540</td>
<td>8</td>
<td>8.6</td>
<td>146</td>
<td>1.44</td>
</tr>
</tbody>
</table>
Implications

More interventions should be conducted as early as possible since less time is wasted to reduce the impact of ASD (Rogers, 1998). In addition, increased social interactions can set the stage for other developments such as generalized use of newly acquired language skills, inclusion in more inclusive educational settings, and development of more positive and long-lasting relationships with peers and other people (Goldstein, Kaczmarek, Pennington, & Shafer, 1992). Results of the present meta-analysis corroborated the belief of the importance of early intervention for children with autism.

More interventions need to be conducted with girls as well as with children diagnosed with Rett Syndrome, CDD, PDD-NOS, and Asperger’s Syndrome. Only 12 studies (21.4%) included 14 girls (11.9%), compared to 44 studies (88.1%) that included 104 boys (78.6%). One hundred and nine children (92.4%) were diagnosed with autism, while three studies included five children (4.2%) with Asperger’s Syndrome, three studies included four children (3.4%) with PDD-NOS, and no study reported any child diagnosed with Rett Syndrome or CDD. The number of girls as the target children was limited, so was the number of children diagnosed with Asperger’s Syndrome or PDD-NOS, partly because of the selection criteria, and partly because of the different prevalence among different genders or different diagnoses.

Future studies need to be conducted using peer-modeling for young children with autism to improve their social interactions. Peer modeling was most effective because it provides visually cued instruction with modeling for children with autism, who demonstrate a preference for visual learning such as the use of visual support instructional strategies (Bellini & Akullian, 2007). Video modeling using peers is one way of peer modeling. This method has the advantage of removing irrelevant stimuli of the modeled behavior through video editing so that the child with autism is able to focus better on the target behavior. In addition, video modeling using peers minimizes human interactions at the stage of acquisition of a new skill and helps reduce the distress and anxiety related to social interactions (Bellini & Akullian). Despite the high efficacy of the peer-modeling intervention, however, only six studies (12.8%) used it.

Results of the analyses indicate that peer-mediated interventions appeared to be moderately effective for improving social initiation. On the other hand, the intervention was highly effective for promoting social responses among young children with ASD. More studies need to be conducted to improve social response since only nine studies studied it (15.3%). Thirteen studies studied social initiation (22%), yet more studies need to be done to investigate how to effectively promote positive social initiations for children with autism.

### TABLE 4 — (Continued)

<table>
<thead>
<tr>
<th>Variables</th>
<th>Frequency</th>
<th>Percent (%)</th>
<th>N</th>
<th>ES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intensity of Intervention (minutes/week)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5–30</td>
<td>13</td>
<td>28.9</td>
<td>671</td>
<td>1.34</td>
</tr>
<tr>
<td>31–60</td>
<td>9</td>
<td>20.0</td>
<td>315</td>
<td>1.06</td>
</tr>
<tr>
<td>61–90</td>
<td>7</td>
<td>15.6</td>
<td>391</td>
<td>0.83</td>
</tr>
<tr>
<td>91–120</td>
<td>3</td>
<td>6.7</td>
<td>78</td>
<td>1.16</td>
</tr>
<tr>
<td>&gt;660</td>
<td>1</td>
<td>2.2</td>
<td>55</td>
<td>1.47</td>
</tr>
<tr>
<td>Not Specified</td>
<td>12</td>
<td>26.7</td>
<td>370</td>
<td>2.76</td>
</tr>
</tbody>
</table>

Note: ** indicates ES is significantly different from 0 ($p < .01$).
* indicates that the effect sizes within the category are significantly different from each other ($p < .05$).
** indicates that the effect sizes within the category are significantly different from each other ($p < .01$).
More studies need to consider maintenance and generalization. Results of this meta-analysis demonstrate that the studies reported maintenance and generalization had higher effect sizes than the studies without reporting maintenance and generalization. In addition, there was a statistically significant difference between the effect sizes. However, the number of studies considering follow-ups and generalization was limited. Eighteen studies reported maintenance (38.3%), 16 studies reported generalization across settings (34.8%), 24 studies reported generalization across behaviors (53.3%), and 21 studies reported generalization across activities (46.7%). The studies considering maintenance or generalization did not occur until the 1990s. The higher effect sizes may be the result of a stronger design and a stronger intervention with maintenance and generalization. However, without the report of the efficacy of follow-ups and generalization, it is unclear how sustainable the intervention is or how feasible and realistic it is to expect school staff, parents/families and other practitioners to carry out interventions once researchers have left (Wheeler, 2007).

Results of this study indicate that the intervention was more effective if the interventionists were siblings. Despite the significance between the effect sizes of peers and siblings, more studies (n = 39) included peers as interventionists (84.8%), compared to five studies (10.9%) which included siblings as interventionists. Results of the present study also suggest the intervention was more effective if it took place at home. Many more studies (n = 28) took place in either integrated or segregated classrooms (57.1%), compared to five studies (10.2%) that took place at home. Additionally, eight studies involved researchers, peers/siblings, school staff, and parents/families (17%), compared to the majority of the studies (n = 37) which excluded parents/families in the intervention (78.7%). Similar to the studies considering maintenance and generalization, the studies with the involvement of all participants did not occur until the 1990s. Researchers have paid attention to the importance of family involvement in the process of intervention for a long time. As early as the 1970s, Lovaas, Koegel, Simmons, and Long (1973) point out that parents of children with autism were critical components of the habilitation process. They demonstrate that it is unlikely to maintain the gains of interventions without parents’ participation (Lovaas et al.). Given the findings of this meta-analysis that the studies were more effective with the involvement of all participants (e.g., the researchers, peers/siblings, school staff, and parents/families) and the importance of families in promoting learning and the value of ongoing interactions with the natural context (Wheeler, 2007), more studies need to be conducted in home and community settings, with siblings as interventionists, and with an active involvement of parents and families.

It is urgent to validate evidence-based practice in the study of social competence among young children with autism given (a) the dramatically increasing prevalence of ASD, (b) social interaction as a core deficit, (c) the heterogeneity of the participants, and (d) varying degree of the educational contexts that serve children with ASD. Results of the meta-analysis indicate the high efficacy of peer-mediated interventions as a method for promoting social interactions among children from birth to eight years of age with ASD. The study also provides integrated findings in detail and thus refines evidence-based peer-mediated intervention practices for young children with ASD. Future studies need to investigate the factors that contribute to more benefits and greater impact for young children with ASD.

References


Hwang, B., & Hughes, C. (2000). The effects of social interactive training on early social commu-


* Sainato, D. M., Goldstein, H., & Strain, P. S.


---

Received: 18 August 2009  
Initial Acceptance: 12 October 2009  
Final Acceptance: 21 November 2009