

# A Meta-Analysis of School-Based Social Skills Interventions for Children With Autism Spectrum Disorders

SCOTT BELLINI, JESSICA K. PETERS, LAUREN BENNER, AND ANDREA HOPF

## ABSTRACT

Social skills deficits are a central feature of autism spectrum disorders (ASD). This meta-analysis of 55 single-subject design studies examined the effectiveness of school-based social skills interventions for children and adolescents with ASD. Intervention, maintenance, and generalization effects were measured by computing the percentage of non-overlapping data points. The results suggest that social skills interventions have been minimally effective for children with ASD. Specific participant, setting, and procedural features that lead to the most effective intervention outcomes are highlighted, and implications for school personnel are discussed. Finally, the results are compared to the outcomes of similar meta-analyses involving social skills interventions with other populations of children.

**I**MPAIRMENT IN SOCIAL FUNCTIONING IS A CENTRAL feature of autism spectrum disorders (ASD) and has been well documented in the literature (Attwood, 1998; Myles et al., 2005; Rogers, 2000). Thus, individuals with ASD have difficulty communicating with others, processing and integrating information from the environment, establishing and sustaining social relationships with others, and participating in new environments. Specific social skills deficits include difficulties with initiating interactions, maintaining reciprocity, sharing enjoyment, taking another person's perspective, and inferring

the interests of others. Although social skills deficits are a central feature of ASD, few children receive adequate social skills programming (Hume, Bellini, & Pratt, 2005). This is a troubling reality, especially considering that the presence of social impairments may portend the development of more detrimental outcomes, such as poor academic achievement, social failure and peer rejection, anxiety, depression, substance abuse, and other forms of psychopathology (Bellini, 2006; La Greca & Lopez, 1998; Tantam, 2000; Welsh, Park, Widaman, & O'Neil, 2001). Most important, social skills deficits impede one's ability to establish meaningful social relationships, which often leads to withdrawal and a life of social isolation. Social skills are critical to successful social, emotional, and cognitive development. As such, effective social skills programming should be an integral component of educational programming for children with ASD.

## SOCIAL SKILLS INTERVENTIONS FOR CHILDREN WITH ASD

A number of qualitative reviews have examined the effectiveness of social skills interventions for children with ASD (Hwang & Hughes, 2000; McConnell, 2002; Rogers, 2000). Hwang and Hughes (2000) reviewed 16 studies involving social skills programming for children with ASD between the

## META-ANALYSES OF SOCIAL SKILLS INTERVENTIONS

ages of 2 and 12 years. The researchers concluded that social skills programming shows “considerable promise for increasing social and communicative skills” (p. 340) of children with ASD, pointing to positive changes in social behaviors across the studies in the literature review. Hwang and Hughes noted that only 9 of 16 studies measured generalization effects, and only two studies assessed intervention fidelity. Rogers (2000) provided a comprehensive narrative of social skills strategies for children with ASD. Similar to Hwang and Hughes, Rogers concluded that children with ASD were responsive to a wide variety of social skills intervention strategies to facilitate both adult-child and child-child interactions. These strategies include pivotal response training, adult prompting, environmental modifications, social skills groups, social stories, video modeling, and peer-mediated instruction.

Based on a literature review of 55 studies examining social skills interventions for young children with ASD, McConnell (2002) concluded that several effective social skills interventions have been documented in the literature and that children with ASD can benefit from social skills programming. Most salient to the present meta-analysis, McConnell divided social skills interventions into five categories: (a) environmental modifications, (b) child-specific interventions, (c) collateral skills interventions, (d) peer-mediated interventions, and (e) comprehensive interventions.

According to McConnell (2002), *environmental modifications* involve modifications to the physical and social environment that promote social interactions between children with ASD and their peers. *Child-specific interventions* involve the direct instruction of social behaviors, such as initiating and responding. *Collateral skills interventions* involve strategies that promote social interactions by delivering training in related skills, such as play behavior and language, rather than training specific social behaviors. *Peer-mediated interventions* involve training nondisabled peers to direct and respond to the social behaviors of children with ASD. Finally, *comprehensive interventions* involve social skills interventions that combine two or more of the aforementioned intervention categories. McConnell’s taxonomy provides a helpful framework for synthesizing studies examining social skills interventions for children with ASD and will be used in the present study.

A limitation of previous literature reviews is the lack of a quantitative metric to evaluate treatment effectiveness. That is, the researchers relied on the conclusions drawn by the studies’ authors to determine treatment effectiveness. Furthermore, such qualitative reviews make it difficult to elucidate the setting, participant, and procedural features that lead to the most beneficial outcomes. Finally, qualitative reviews do not allow relative comparisons of treatment effectiveness across different intervention strategies. A quantitative synthesis of studies examining social skills interventions would add a substantial amount of information to our existing knowledge regarding the effectiveness of these interventions for children with ASD.

A number of quantitative meta-analyses have been performed on social skills intervention studies involving populations of children and adolescents other than persons with ASD. In general, these studies have demonstrated that traditional social skills training programs are only minimally effective in teaching social skills to children and adolescents (Gresham, Sugai, & Horner, 2001; Quinn, Kavale, Mathur, Rutherford, & Forness, 1999). Quinn et al. (1999) found small effect sizes in their meta-analysis of 35 studies examining social skills interventions in children and adolescents with emotional and behavioral disorders. Moreover, no significant differences in outcomes were observed for duration of the intervention, quality of the research design, age of the participant, and the specific construct used to measure social skills. The researchers concluded that social skills programs must be designed to fit the individual needs of the child, as opposed to forcing the child to “fit” into the chosen social skills strategy or strategies. Furthermore, the researchers concluded that the *type* of skill deficit (i.e., performance deficit versus skill acquisition deficit) must also be considered when developing a social skills intervention plan. A *performance deficit* refers to a skill or behavior that is present but not demonstrated or performed, whereas a *skill acquisition deficit* refers to the absence of a particular skill or behavior.

Mathur, Kavale, Quinn, Forness, and Rutherford (1998) conducted a meta-analysis of 64 single-subject research studies involving social skills training for children with emotional and behavioral disorders, including children with ASD. The specific focus was on treatment effects of social skills training based on computing the *percentage of non-overlapping data points* (PND) between the baseline and treatment phases for each of the studies reviewed. The overall mean treatment PND of 62% indicated a low treatment effect. The results of this study were nearly identical to those of the meta-analysis conducted by Scruggs and Mastropieri (1994), who also found a mean treatment PND of 62% for social skills training. Most salient to the present study, the mean treatment PND for the subsample of children with ASD was lower ( $M = 54%$ ) than for the subsample of children with emotional and behavioral disorders ( $M = 64%$ ) and the subsample of children with delinquent behaviors ( $M = 76%$ ). Although the mean generalization PND was also low, the mean maintenance PND of 74% indicated a moderate maintenance effect for social skills training. Furthermore, the meta-analysis revealed that the lowest treatment effects were in the preschool group compared to elementary and secondary students. The researchers concluded that this age effect might be attributed to the fact that many social skills interventions are not structured to meet the needs of preschoolers. The meta-analysis also revealed that the number of treatment sessions was not associated with treatment outcomes. Many of the studies provided vague descriptive information regarding the duration of

instruction, however; therefore, these results should be interpreted with caution.

After reviewing numerous studies, Gresham et al. (2001) concluded that meta-analytic reviews of social skills training have yielded a wide variety of results, ranging from ineffectual to highly effective interventions. Based on their review of the literature, Gresham et al. provided a number of recommendations for promoting effective social skills interventions. First, they recommended that social skills training should be implemented more frequently and more intensely than what is typically implemented. They concluded that “thirty hours of instruction, spread over 10–12 weeks is not enough” (p. 341). Second, they concluded that a major weakness of social skills interventions is a failure to produce adequate maintenance and generalization effects. Gresham et al. attributed this, in part, to the fact that social skills training often takes place in “contrived, restricted, and decontextualized” (p. 340) settings, such as resource rooms or other pullout settings. Third, the researchers posited that the ineffectiveness of many social skills programs is a result of the interventionists’ failure to match the social skills strategy to the type of skill deficit presented. For instance, if the child is experiencing skill acquisition deficits, then intervention strategies designed to teach new skills should be selected. If the child is experiencing performance deficits, then interventions strategies designed to enhance the performance of existing skills should be selected. Finally, Gresham et al. found that few meta-analytic studies reported evidence that the interventions were implemented as intended. This absence of fidelity data makes it extremely difficult to conclude whether a social skills intervention was ineffective due to an ineffectual intervention strategy or because it was poorly implemented.

## SOCIAL SKILLS INTERVENTIONS IN SCHOOLS

The school setting provides a fertile ground for delivering effective social skills programming, but it also presents formidable obstacles. The school day is filled with abundant opportunities to interact with peers in a natural social environment. Also, schools are equipped with trained professionals who are qualified to teach social skills, such as teachers, therapists, and psychologists. Schools are often relied upon to shoulder the responsibility of delivering social skills programming to children with social skills deficits, because the presence of these deficits significantly interferes with social relationships and has a deleterious impact on academic performance (Welsh et al., 2001). However, implementing social skills programming in schools can be challenging for school personnel, who often are presented with limited time, resources, and training. Despite a need to identify and deliver effective school-based social skills training to children with ASD, no studies to date have examined the effectiveness of this type of training for this particular population. The present

study provides a meta-analysis of school-based social skills intervention studies involving children with ASD.

## PURPOSE OF STUDY

The purpose of this study was to provide a quantitative synthesis of existing single-subject research studies on school-based social skills interventions for children with ASD. Another purpose was to examine the aggregated outcomes of these studies and to identify the participant, setting, and procedural features that lead to the most effective intervention outcomes for children with ASD. The final purpose was to compare the intervention, maintenance, and generalization effects of the studies to the outcomes of similar studies involving social skills interventions with other populations of children.

## METHOD

### Locating Studies

Studies included in the meta-analysis were located by conducting a search of journal articles using the Educational Resources Information Center (ERIC) and PsycINFO databases for the years 1980–2005. Searches were carried out using a combination of the following descriptors: *autism, autism spectrum disorder, ASD, pervasive developmental disorder, PDD, Asperger’s, Asperger syndrome, social skills, social group, social interaction, social behavior, conversation, compliments, cooperation, social communication, social initiation, social response, play, eye contact, turn-taking, and reciprocity*. Furthermore, we conducted an ancestral search of studies using the reference lists of each study located via ERIC and PsycINFO in an effort to locate additional studies that were not captured by the initial database search. Finally, manual searches of the journals *Focus on Autism and Other Developmental Disabilities* and *Journal of Autism and Developmental Disorders* were conducted. In all, 514 studies were located in this initial search of journal articles.

The abstract or method section of each article was reviewed to determine inclusion eligibility based on the following nine criteria. First, participants must have been identified as having ASD. In some instances, studies included both children with ASD and children without ASD. In these cases, data were analyzed only for the participants with ASD. Second, the study must have used outcome measures that targeted social functioning. Studies that measured only academic functioning or the reduction of problem behaviors were excluded. Third, the study must have assessed the efficacy of social skills interventions. Interventions that measured social behaviors but did not use a social skills intervention (e.g., pharmacological or other medical interventions) were excluded. Fourth, the study must have examined social skills interventions that were implemented in a school setting. In

cases where studies involved multiple settings (e.g., home and school), only the data for the interventions that took place in the school setting were analyzed. Fifth, the study must have used a single-subject research design. Sixth, studies must include dichotomous dependent variables (e.g., *yes-no*, *correct-incorrect*) with at least three probes or questions per data point or three data points per intervention phase. The reason for this criterion was that fewer instances could not be logically or intuitively interpreted by the metric employed in the meta-analysis (i.e., PND). Seventh, the study must present data in graphical displays that depicted individual data points rather than aggregated data (such as means), as these graphical displays were critical to the determination of intervention effectiveness. Eighth, the study must have been published in a peer-reviewed journal. Finally, the study must have been published in English. In all, 55 studies met all these selection criteria and were included in the quantitative analysis.

## Classification

A coding system, adapted from Bellini and Akullian (2007), was designed to analyze the 55 studies. Each study was analyzed across the following categories:

1. participant characteristics, including number of participants, diagnosis, and age;
2. intervention description, including the type of intervention strategy employed and the format of the intervention (e.g., individual versus group training);
3. length and dosage of intervention, including number of treatment sessions, hours of intervention, and duration of intervention (number of days from the beginning of the intervention to the end of the intervention);
4. research design;
5. description of the targeted skills and dependent variables;
6. intervention effectiveness, including intervention, maintenance, and generalization effects as measured by the percentage of non-overlapping data points;
7. location of intervention (e.g., classroom or resource/therapy room);
8. confirmation of experimental control through the introduction or withdrawal of the independent variable across three points in time or across three data series (Horner et al., 2005);
9. confirmation of whether the study measured interobserver reliability, intervention fidelity, and social validity; and

10. confirmation of whether the study systematically matched the intervention strategy to the type of skill deficit.

Three authors independently coded each study. To ensure interrater reliability, the primary author independently coded all 55 studies. Interrater agreement was obtained by dividing the total number of agreements by the total number of agreements plus disagreements and multiplying by 100. Interrater agreement for study features was 94%, with a range from 89% to 100%. Given the fact that PND calculations are more complex than recording study features, a separate procedure was employed to ensure the accuracy of the PND calculations. PND scores were initially calculated by two of the authors. To ensure interrater reliability, the primary author also calculated PND scores for each study. Initial interrater agreement for PND calculations was 81%, with a range from 71% to 91%. Discrepancies were resolved through discussion and further inspection of the graphed data points. PND scores were then independently recalculated by the two authors who initially calculated PND scores, resulting in an interobserver agreement of 100%. Furthermore, one independent reviewer (a graduate student trained in the coding system and PND calculations) coded 14 randomly selected studies (25% of the studies used in the meta-analysis) using the same procedures. The mean interrater agreement between the independent reviewer and the authors was 100% for study features and 93% for PND calculations.

## Analysis

The effectiveness of interventions was determined by computing the PND for each study. PND provides a measure of intervention effectiveness and a method for systematically synthesizing single-subject research studies (Mastropieri & Scruggs, 1985–1986). It is determined by calculating the percentage of intervention data points that do not overlap with the highest baseline data point. Scruggs and Mastropieri (2001) argued that the use of PND is preferable to the use of a conventional effect size (*ES*) in synthesizing single-subject research for two primary reasons. First, *ES* computations are derived theoretically from procedures used in inferential statistics. This is problematic because the data derived from single-subject research are not independent, thereby violating a primary assumption of inferential statistics. Second, many single-subject studies include relatively few data points, which may inflate the *ES*, thus making interpretation difficult at best. According to Scruggs and Mastropieri (1998), PND scores above 90 represent very effective intervention scores, scores from 70 to 90 represent effective interventions, scores from 50 to 70 represent interventions with low or questionable effectiveness, and scores below 50 represent ineffective interventions.

PND scores were calculated for each participant and across all dependent variables measured in each of the 55 studies. Specifically, PND scores were calculated for inter-

vention effects, maintenance effects, and generalization effects. Maintenance effects were measured by calculating PND between baseline and the maintenance phase. Generalization effects were calculated for studies that measured intervention effectiveness across persons, settings, play stimuli, or skills and behaviors. Mean PND scores were calculated for each study and aggregated for the entire data set. Mean PND scores are reported across the two categories of dependent variables included in the analysis: collateral skills and social interaction skills. Collateral skills (McConnell, 2002) included play skills, language/conversation skills, prosocial skills (e.g., sharing, waiting for a turn, social smiling), eye contact, and joint attention. Social interaction skills included specific social behaviors, such as social initiations, social responses, social participation, and duration of interactions. One study measured both categories of dependent variables. In this case, mean PND scores were calculated for the overall study and for each of the outcome variables. PND scores are also reported across intervention features, including (a) type of intervention, as defined by McConnell (2002; e.g., environmental modifications, collateral skills intervention, peer-mediated intervention, child-specific intervention, or comprehensive intervention, which consists of a combination of the other strategies); (b) length and duration of intervention; (c) intervention format (e.g., individual versus group training); and (d) location of intervention (e.g., typical classroom or pullout). For the latter analysis, an intervention that involved removing the child from his or her typical classroom setting for any portion of the intervention was coded as a pullout intervention. One study used classroom interventions for four participants but used a pull-out intervention for one participant. In this case, mean PND scores were calculated for the overall study, and separate mean PND scores were also calculated based on the location of the intervention. Finally, PND scores are reported across participant grade level (e.g., preschool, elementary, and secondary). In some studies, the authors did not indicate grade level for elementary and secondary level children. In these cases, children under 13 years old were coded as elementary, whereas children 13 years and older were coded as secondary. Six studies involved children from multiple age groups. In these cases, mean PND scores were calculated for the overall study and for each of the age groups represented in the study.

Given the nonparametric nature of PND, the Kruskal-Wallis procedure—a nonparametric test that allows for the comparison of multiple independent samples—was performed to test for significant differences in PND across dependent variables (collateral skills vs. social interaction skills), intervention features (type of intervention, intervention format, and location of intervention), and age groups (preschool, elementary, and secondary). Furthermore, Pearson product-moment correlations were conducted to examine relationships between length and duration of intervention and treatment outcomes. Statistical significance was established at the .05 level.

## Overall Study Characteristics

The 55 peer-reviewed studies included in the meta-analysis were published between 1986 and 2005. A total of 157 participants were included. Forty-two studies used a variation of a multiple baseline or probe design, six studies used a reversal design, three studies used an A–B design, two studies used a changing conditions design, one study used an alternating treatment design, and one study used an alternating treatment and reversal design. Of the 42 studies using a multiple baseline or probe design, 24 used a multiple baseline–only design, and 18 used a combination of multiple baseline and changing conditions, reversal, or alternating treatment design. Interobserver reliability was reported in all 55 studies. Social validity was measured in 12 studies, and intervention fidelity was measured in 14 studies. Experimental control was demonstrated in 49 studies through the introduction or removal of the independent variable across three or more points in time. Interventions ranged in length from 8 to 73 sessions ( $n = 38$ ,  $Mdn = 25.5$  sessions). Hours of intervention ranged from 2.5 to 28 hours ( $n = 18$ ,  $Mdn = 7.25$  hours). Length of intervention ranged from 10 to 210 days ( $n = 30$ ,  $Mdn = 60$  days). Finally, only one study systematically matched the type of intervention strategy with the type of skill deficit exhibited by participants.

## Intervention, Maintenance, and Generalization Effects

Based on Scruggs and Mastropieri's (1998) recommendations for interpreting the magnitude of PND scores, as noted earlier, the results indicated that social skills interventions produce low to questionable treatment and generalization effects and moderate maintenance effects for children with ASD. The results indicated a questionable intervention effect for the studies in the meta-analysis ( $n = 52$ ; PND  $M = 70\%$ , range = 17–100%). Just over half of the reviewed studies (52%) produced mean intervention PND scores above 70. Three studies did not provide individual data points for the intervention phase of the studies; therefore, intervention PND was not calculated for these studies. Maintenance effects were reported in 25 studies, and generalization was reported and graphed in 15 studies. The results indicated moderate maintenance effects ( $n = 25$ ; PND  $M = 80\%$ , range = 17–100%) and low generalization effects ( $n = 15$ ; PND  $M = 53\%$ , range = 17–100%).

## Differences Across Intervention and Participant Features

Table 1 presents mean PND scores for intervention, maintenance, and generalization effects across the types of intervention used in the studies. As illustrated, 10 studies used a

collateral skills intervention, 10 used a peer-mediated intervention, 15 used a child-specific intervention, and 20 used a comprehensive intervention. No studies used an exclusively environmental modification intervention. Intervention effects were similar across categories, with the highest PND scores observed for collateral skills interventions (range = 43–100%), and the lowest scores observed for peer-mediated interventions (range = 37–100%). Greater variation in mean PND scores was found for maintenance effects, with the highest scores observed for collateral skills interventions (range = 88–100%) and the lowest scores for comprehensive interventions (range = 35–100%). Generalization PND scores also varied considerably and were likely influenced by the small number of studies that measured generalization effects. The Kruskal-Wallis procedure was performed to test for significant differences in intervention, maintenance, and generalization effects across type of social skills intervention. None of these differences reached the .05 level of significance.

Table 2 presents mean PND scores for intervention, maintenance, and generalization effects for individual and group interventions. As illustrated, 24 studies investigated interventions that involved only the target child, whereas 31 used a group intervention. Low to moderate intervention effects were found for both individual (range = 17–96%) and group interventions (range = 37–100%). Moderate maintenance effects were observed for both intervention formats. Higher generalization effects were found for individual inter-

ventions than for group interventions. Results of the Kruskal-Wallis procedure revealed that the differences between the groups were not statistically significant.

Table 3 presents mean PND scores for classroom and pullout interventions; 27 studies delivered the social skills intervention within the child's typical classroom setting, and 20 implemented a pullout intervention. One study used a classroom intervention for four participants but used a pullout intervention for one student. Seven studies did not provide sufficient information to discern the location of the intervention. The Kruskal-Willis procedure indicated significant differences between the two groups across intervention,  $\chi^2(1) = 4.63, p < .05$ ; maintenance,  $\chi^2(1) = 4.08, p < .05$ ; and generalization effects,  $\chi^2(1) = 5.52, p < .05$ . Interventions delivered in the child's typical classroom setting produced moderate intervention and maintenance effects and low generalization effects, whereas pullout interventions produced low to questionable intervention and maintenance effects and very low generalization effects.

Table 4 presents mean PND scores across age groups (preschool, elementary, and secondary). As illustrated, 21 studies involved preschool-age children, 23 involved elementary-age children, and 5 studies involved secondary-age students. Six studies involved participants from multiple age groups. Intervention effects were similar across age groups, with the highest PND scores for secondary-age children (range = 21–98%) and the lowest PND scores observed for elementary-

**TABLE 1. Mean PND Scores by Type of Social Skills Intervention**

Intervention type	Intervention		Maintenance		Generalization	
	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>
Collateral skills	75	7	93	5	51	2
Peer mediated	62	10	79	5	33	3
Child specific	71	15	87	5	67	1
Comprehensive	72	20	69	8	59	9

Note. PND = percentage of non-overlapping data points.

**TABLE 2. Mean PND Scores by Format of Social Skills Intervention**

Format	Intervention		Maintenance		Generalization	
	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>
Individual	72	24	84	12	63	8
Group	69	28	77	13	43	7

Note. PND = percentage of non-overlapping data points.

**TABLE 3. Mean PND Scores by Location of Social Skills Intervention**

Location	Intervention		Maintenance		Generalization	
	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>
Classroom	76	28	88	16	67	8
Pullout	62	21	67	9	29	5

Note. PND = percentage of non-overlapping data points.

**TABLE 4. Mean PND Scores by Age Group**

Age group	Intervention		Maintenance		Generalization	
	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>
Preschool	70	21	74	8	65	7
Elementary	69	28	79	14	40	7
Secondary	76	9	99	5	71	2

Note. PND = percentage of non-overlapping data points.

age children (range = 17–100%). The highest maintenance effects were noted for the secondary group (range = 96–100%) and the lowest for the preschool group (range = 17–100%). The highest generalization effects were also observed for the secondary group (range = 67–75%), with the lowest PND scores for elementary-age children (range = 16–79%). The Kruskal-Wallis procedure was performed to test for significant differences across age groups. None of the differences reached the .05 level of significance.

Correlational analyses were conducted to examine relationships among the number of intervention sessions, hours of intervention, and length of treatment sessions and treatment outcomes (intervention, maintenance, and generalization effects). The strength of these associations was low, and none of the correlations reached the .05 level of significance.

### Differences Across Dependent Variables

PND scores were calculated for two categories of dependent variables: collateral skills and specific social behaviors. Thirty-six studies targeted specific social behaviors, and 18 studies targeted collateral skills. One study measured both types of dependent variables. Similar intervention, maintenance, and generalization effects were observed between collateral skills and specific social behaviors. Interventions targeting collateral skills produced low to moderate intervention ( $n = 17$ ; PND  $M = 72%$ , range = 25–100%), maintenance ( $n = 11$ ; PND  $M = 75%$ , range = 33–100%), and generalization effects

( $n = 6$ ; PND  $M = 53%$ , range = 25–80%). Interventions targeting specific social behaviors also produced low to moderate intervention ( $n = 36$ ; PND  $M = 69%$ , range = 17–97%), maintenance ( $n = 15$ ; PND  $M = 85%$ , range = 35–100%), and generalization effects ( $n = 11$ ; PND  $M = 56%$ , range = 16–84%). The Kruskal-Wallis procedure revealed no statistically significant differences between dependent variables.

## DISCUSSION

The results of this meta-analysis suggest that school-based social skills interventions are minimally effective for children with ASD. Specifically, social skills interventions produced low treatment effects and low generalization effects across participants, settings, and play stimuli. Moderate maintenance effects were observed, suggesting that gains made via social skills interventions are maintained after the intervention is withdrawn. Furthermore, similar intervention, maintenance, and generalization effects were observed between interventions targeting collateral skills (e.g., play skills, joint attention, and language skills) and interventions targeting specific social behaviors (e.g., social initiations, social responses, and duration of interaction).

The low treatment effects observed in the present study are consistent with the results of previous social skills intervention meta-analyses (Mathur et al., 1998; Scruggs & Mastropieri, 1994). The mean PND scores for intervention effects

observed in the present study ( $M = 70\%$ ) were higher than the mean PND scores observed for the subsample of participants with ASD in Mathur et al.'s study ( $M = 54\%$ ). This difference may be due to the fact that the present study included only school-based interventions, whereas Mathur et al. included interventions that were implemented in multiple settings. The differences in outcomes may also be attributed to differences in sample size. Mathur et al. did not report the total number of participants with ASD, nor did they report the total number of studies involving participants with ASD. However, the ASD group was one of three groups derived from a total sample of 64 studies and 283 participants. In comparison, the present study included 55 studies, involving 157 participants with ASD. It can be reasonably inferred that the subsample of participants with ASD included in Mathur et al.'s meta-analysis was considerably smaller than the sample of participants with ASD included in the present meta-analysis.

Similar treatment, maintenance, and generalization effects were observed across collateral skills, peer-mediated, child-specific, and comprehensive interventions. Furthermore, there were no significant differences between the intervention and maintenance effects of studies that implemented group interventions and studies that implemented individual interventions. Individual interventions produced higher generalization effects than did group interventions; however, these differences were not statistically significant. An examination of outcomes across age groups revealed that social skills interventions were most effective with secondary-age students. However, the differences across the groups were not significant and should be interpreted with caution given the small number of studies ( $n = 9$ ) that involved secondary-age participants. The length and duration of interventions varied considerably across studies. There were no significant relationships observed between the number of intervention sessions, hours of intervention, and length of intervention (total days from the beginning to the end of the intervention) and treatment outcomes. However, statistically significant differences were observed between interventions implemented in the child's typical classroom and studies implemented in pullout settings. That is, studies implemented in the child's typical classroom setting produced significantly higher intervention, maintenance, and generalization effects than interventions that involved removing the child from the classroom. Finally, only one study systematically matched the type of intervention strategy with the type of skill deficits exhibited by participants.

## Implications for Practice

As mentioned earlier, Gresham et al. (2001) provided several recommendations for improving the effectiveness of social skills interventions. Their recommendations included increasing the dosage of intervention, providing instruction in the child's natural setting, matching the intervention strategy with the type of skill deficit, and ensuring intervention fi-

delity. These recommendations are discussed in relation to the findings of the present meta-analysis, and implications for both research and practice are provided.

**Dosage.** Gresham et al. (2001) recommended that social skills interventions be implemented more intensely and frequently than the level presently delivered to children with social skills deficits. Although Gresham et al. did not recommend a specific dosage, they stated that 30 hours of instruction, spread over 10 to 12 weeks, was insufficient. The low intervention effects observed in the present meta-analysis may be attributed to the low level of instructional intensity provided in the reviewed studies, which was considerably lower than the 30-plus hours recommended. It is important to note that the present study found no significant relationships between length of intervention and intervention outcomes. However, as in Mathur et al. (1998), many studies failed to provide adequate descriptive information regarding the intervention length and duration. Therefore, the results related to intervention length and study outcomes should be interpreted with caution. Children with ASD exhibit significant social skills deficits that may potentially lead to academic, behavioral, and emotional difficulties. As such, the recommendation to increase instructional intensity is particularly salient for this population of children. School personnel should look for opportunities to teach and reinforce social skills as frequently as possible throughout the school day. Future research studies should provide more descriptive information regarding intervention length and examine the relationship between length of intervention and intervention outcomes.

**Intervention Setting.** Gresham et al. (2001) noted that the weak outcomes of social skills interventions can be attributed to the fact that these interventions often take place in "contrived, restricted, and decontextualized" (p. 340) settings, such as resource rooms or other pullout settings. According to Gresham et al., this "decontextualized" programming leads to poor maintenance and generalization effects. The results of the present meta-analysis support this assertion; that is, maintenance and generalization effects were significantly lower for interventions that were implemented in pullout settings. In contrast, interventions that were implemented in the child's typical classroom setting produced higher maintenance effects and higher generalization effects across participants, settings, and play stimuli. In addition to higher maintenance and generalization effects, the results of the present study also suggest that interventions implemented in the child's typical classroom produce higher intervention effects. This finding has important implications for school-based social skills interventions. Teachers and other school personnel should place a premium on selecting social skills interventions that can be reasonably implemented in multiple naturalistic settings. This is particularly important for children with ASD, who may have considerable difficulties with transferring skills from one setting to another. Future research

should be conducted to examine this finding more comprehensively. For instance, comparison studies can be conducted to examine the differential effects of identical interventions implemented in multiple settings.

**Matching Strategy with Type of Skill Deficit.** Gresham et al. (2001) asserted that a key component of effective social skills programming is the ability of the interventionist to match the intervention strategy with the type of skill deficit. This position was supported by Quinn et al. (1999), who concluded that the failure of many social skills interventions results from a mismatch between strategy and skill deficits. Quinn et al. further stated that interventions should be designed to address the individual needs of the child, rather than forcing the child to “fit” into the selected strategy. Of the 55 studies included in the present meta-analysis, only one study pointed out the type of skill deficit exhibited by the participants. School personnel should make an intensive effort to systematically match the intervention strategy to the type of skill deficit exhibited by the child. For instance, if the child lacks the skills necessary to join in an interaction with peers, an intervention strategy should be selected that promotes skill acquisition. In contrast, if the child has the skills to join in an activity but regularly fails to do so, a strategy should be selected that enhances the performance of the existing skill. See Bellini (2006) for a comprehensive discussion of how to discern between a skill acquisition and a performance deficit and for a list of social skills intervention strategies that have been divided into strategies that promote skill acquisition and strategies that enhance the performance of existing skills.

**Intervention Fidelity.** Only 14 of the studies in the present meta-analysis measured whether the intervention was implemented as intended. Gresham et al. (2001) concluded that the failure of studies to provide intervention fidelity data makes it extremely difficult to conclude whether a social skills intervention was ineffective because of an ineffectual intervention strategy or because the strategy was poorly implemented. Poor intervention fidelity may significantly diminish the outcomes of the social skills intervention and diminish the implementers’ ability to make decisions regarding the effects of individual strategies. Furthermore, social validity data were collected in only 12 of the studies in the analysis. *Social validity* refers to the social significance of the treatment objectives, the social significance of the intervention strategies, and the social importance of the intervention results (Gresham & Lambros, 1998). Thus, it involves ensuring that the consumers (parents and teachers) believe that the selected intervention strategies are effective and appropriate, and that the social objectives are important for the child to achieve. Social validity often has a direct impact on intervention fidelity. If the intervention lacks social validity, consumers are less likely to exert the effort necessary to implement the intervention, thus diminishing intervention fi-

delity. School personnel and researchers should make a concerted effort to collect data related to intervention fidelity.

## Limitations and Suggestions for Future Research

Although the results of the present study have valuable implications for school personnel and researchers, several limitations must be noted. The relatively small number of studies included in the meta-analysis precluded a thorough analysis of covariation between participant characteristics (age-specific diagnosis, cognitive level, language level), setting characteristics (classroom vs. pullout), intervention features (length of intervention, type of intervention, group vs. individual intervention), and the outcomes associated with the social skills interventions. In particular, the sample size precluded examining interaction effects among intervention features, participant features, and outcomes. For instance, it would be important to examine whether classroom interventions are more or less effective for secondary-age children than they are for preschool-age children or elementary-age children. The number of studies included in the meta-analysis also precluded us from comparing the outcomes of specific social skills strategies, such as social stories, social scripting, prompting, video modeling, and so on. Future research should examine the outcomes of specific strategies and elucidate the participant, setting, and procedural features that lead to the most beneficial social outcomes for children with ASD.

It is important to interpret the outcomes of the present meta-analysis in the context of a potential publication bias. *Publication bias* refers to the tendency of researchers and journal editors to make publication decisions based on the strength of the study findings (Scargle, 2000). Thus, published intervention studies may inflate the true effects of a particular intervention. Publication bias presents a particular obstacle to meta-analytical researchers, as it could potentially lead researchers to overestimate the positive effects of a particular intervention. In the case of the present meta-analysis, publication bias may have inflated the intervention, maintenance, and generalization effects of the school-based social skills interventions. If so, this would indicate that the true effects of social skills interventions for children with ASD may be even lower than the minimal effectiveness observed in the present meta-analysis. Future research should examine the threat of publication bias by including unpublished studies, such as dissertations, in the meta-analysis and comparing their results to those of published studies.

Only 15 studies measured the generalization effects of social skills interventions. This small subsample precluded a more comprehensive examination of intervention features that promote the generalization of skills across participants and settings. Generalization of skills is an integral component of social skills interventions, and the success of social skills interventions should be judged based on the ability of the child to perform the skill in multiple social settings and with multiple persons. School personnel should create an explicit plan

for promoting generalization when developing a social skills intervention. Future research should elucidate features that improve generalization effects, such as programming for generalization, teaching self-monitoring, or providing instruction in the natural setting.

The present meta-analysis only included school-based social skills interventions. It is unclear how the results of these school-based interventions compare to those of interventions implemented in other settings, such as clinical or home settings. Future research using similar methodology should examine the outcomes of social skills interventions implemented in these other settings and compare these results with the outcomes reported in the present study.

## Summary

The results of the present study are consistent with those of previous meta-analyses indicating that social skills interventions are only minimally effective for children with ASD and other social skills deficits. The results support the recommendations offered by Gresham et al. (2001), which include increasing the dosage of social skills interventions, providing instruction in the child's natural setting, matching the intervention strategy with the type of skill deficit, and ensuring intervention fidelity. Future research is needed to examine both the efficacy and the social validity of social skills interventions and to examine factors that lead to the most beneficial outcomes for children with ASD. ■

**SCOTT BELLINI**, PhD, is the assistant director of the Indiana Resource Center for Autism and a faculty member with the Indiana University school psychology program. His research and clinical interests include social-emotional functioning for children with autism spectrum disorders. **JESSICA K. PETERS**, BSc, is a doctoral student in the school psychology program at Indiana University. She is interested in autism spectrum disorders, early intervention, and children who speak English as a second language. **LAUREN BENNER**, BA, is a doctoral student in the school psychology program at Indiana University. Her research interests include early diagnosis and intervention for children with autism spectrum disorders. **ANDREA HOPF**, MSc, is a doctoral student in the school psychology program at Indiana University. Her interests include educational programming for students with autism spectrum disorders and academic and behavioral interventions. Address: Scott Bellini, PhD, Indiana Resource Center for Autism, 2853 E. 10th St., Bloomington, IN 47408; e-mail: sbellini@indiana.edu

## REFERENCES

- Attwood, T. (1998). *Asperger's syndrome: A guide for parents and professionals*. Philadelphia: Kingsley.
- Bellini, S. (2006). The development of social anxiety in high functioning adolescents with autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities, 21*, 138–145.
- Bellini, S., & Akullian, J. (2007). A meta-analysis of video modeling and video self-modeling interventions for children and adolescents with autism spectrum disorders. *Exceptional Children, 73*, 261–284.
- Gresham, F. M., & Lambros, K. M. (1998). Behavioral and functional assessment. In T. S. Watson & F. M. Gresham (Eds.), *Handbook of child behavior therapy* (pp. 3–22). New York: Plenum Press.
- Gresham, F. M., Sugai, G., & Horner, R. H. (2001). Interpreting outcomes of social skills training for students with high-incidence disabilities. *Teaching Exceptional Children, 67*, 331–344.
- Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, A., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Council for Exceptional Children, 2*, 165–179.
- Hume, K., Bellini, S., & Pratt, C. (2005). The usage and perceived outcomes of early intervention and early childhood programs for young children with autism spectrum disorder. *Topics in Early Childhood Special Education, 25*, 195–207.
- Hwang, B., & Hughes, C. (2000). The effects of social interactive training on early social communicative skills of children with autism. *Journal of Autism and Developmental Disorders, 30*, 331–343.
- La Greca, A. M., & Lopez, N. (1998). Social anxiety among adolescents: Linkages with peer relations and friendships. *Journal of Clinical Child Psychology, 26*, 83–94.
- Mastropieri, M. A., & Scruggs, T. E. (1985–1986). Early intervention for socially withdrawn children. *The Journal of Special Education, 19*, 429–441.
- Mathur, S. R., Kavale, K. A., Quinn, M. M., Forness, S. R., & Rutherford, R. B. (1998). Social skills interventions with students with emotional and behavioral problems: A quantitative synthesis of single subject research. *Behavioral Disorders, 23*, 193–201.
- 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 research. *Journal of Autism and Developmental Disorders, 32*, 351–372.
- Myles, B. S., Adreon, D. A., Hagen, K., Holverstott, J., Hubbard, A., Smith, S. M., et al. (2005). *Life journey through autism: An educator's guide to Asperger syndrome*. Arlington, VA: Organization for Autism Research.
- Quinn, M. M., Kavale, K. A., Mathur, S. R., Rutherford, R. B., Jr., & Forness, S. R. (1999). A meta-analysis of social skills interventions for students with emotional and behavioral disorders. *Journal of Emotional and Behavioral Disorders, 7*, 54–64.
- Rogers, S. (2000). Interventions that facilitate socialization in children with autism. *Journal of Autism and Developmental Disorders, 30*, 399–409.
- Scargle, J. D. (2000). Publication bias: The “file-drawer” problem in scientific interference. *Journal of Scientific Exploration, 14*, 91–106.
- Scruggs, E. T., & Mastropieri, M. A. (1994). The effectiveness of generalization training: A quantitative synthesis of single subject research. In T. E. Scruggs & M. A. Mastropieri (Eds.), *Advances in learning and behavioral disabilities* (Vol. 8, pp. 259–280). Greenwich, CT: JAI Press.
- Scruggs, E. T., & Mastropieri, M. A. (1998). Summarizing single-subject research: Issues and applications. *Behavior Modification, 22*, 221–242.
- Scruggs, E. T., & Mastropieri, M. A. (2001). How to summarize single-participant research: Ideas and applications. *Exceptionality, 9*, 227–244.
- Tantam, D. (2000). Psychological disorder in adolescents and adults with Asperger syndrome. *Autism, 4*, 47–62.
- Welsh, M., Park, R. D., Widaman, K., & O'Neil, R. (2001). Linkages between children's social and academic competence: A longitudinal analysis. *Journal of School Psychology, 39*, 463–481.