EBP
briefs

A scholarly forum for guiding evidence-based
practices in speech-language pathology

Improving Communication for Children with Autism:
Does Sign Language Work?

Jamie B. Schwartz, PhD
Chad Nye, PhD

EBP Briefs

A publication of AGS Publishing, now part of Pearson Assessements

Improving Communication for Children with Autism:
Does Sign Language Work?

Jamie B. Schwartz, PhD
Chad Nye, PhD
University of Central Florida

Improving Communication for Children with Autism: Does Sign Language Work?

Jamie B. Schwartz, PhD
Chad Nye, PhD
University of Central Florida

One of the signature characteristics of children with autism is failure to develop adequate communication skills. Despite efforts to teach oral communication skills to children with autism, many of these children continue to experience difficulties in acquiring functional speech. Therefore, clinicians often are faced with the decision of selecting and implementing an aided or unaided augmentative or alternative communication system for these individuals. Although aided communication systems, ranging from low (e.g., pictures) to high tech (e.g., speech out put devices), have been used with children with autism [see Miranda (2002) for a review], unaided systems of communication (e.g., gestural communication or sign language) continue to be recommended as well to provide children with autism an augmentative or alternative means of communication. Given that a clinician may recommend sign language training for a child with autism, what evidence is available upon which to base this decision? At the heart of evidence–based practice is the integration of the best research evidence in the decision making process. This paper will investigate the evidence available to characterize whether teaching sign language [i.e., sign language alone or total communication (sign + speech)] will improve either the sign or oral communication of children with autism.

Background Literature

A number of summary review papers are available that discuss the effectiveness of intervention programs for teaching sign language to children with autism (Creekmore, 1982; Goldstein, 2002; Kiernan, 1983). Each of these reviews evaluated evidence in which the intervention included sign language alone or in combination with speech (i.e., total communication),

and each concluded that a case could be made for use of either sign language or total communication to facilitate the communication competence of children with autism. However, the evidence upon which the conclusions were drawn was based on research using single subject designs. Kiernan (1983) pointed out that single subject design studies provide valuable data and, for well designed studies, can support the effectiveness of treatments for participants in the study, although the results of these studies cannot be generalized readily to the populations represented by the individuals in a study. What also is not clear from these reviews is just how much improvement in communication one might expect as a result of teaching children with autism sign language or total communication.

Clinicians seeking to implement evidence-based practices for teaching children with autism might ask questions such as: (1) Does existing research provide a data-based comparison of the effects of interventions using sign language or total communication to improve the communication skills of children with autism? And (2) Can an evidence-based practice recommendation be advanced based on these data? These types of questions need to be based on a more rigorous and comprehensive quantitative evaluation of the research than has previously occurred, namely through the use of systematic review and meta-analysis. The purpose of this paper is to use these tools to summarize and synthesize existing research examining the efficacy of sign language intervention (sign alone or total communication) to improve the sign or oral communication skills of children with autism.

Research summaries and meta-analyses should guide evidence-based practice.

Method

This research used a retrospective procedure known as systematic review and meta-analysis to provide an interpretation of the research data which clinicians can use when making an evidence-based best practice decision. The use of a retrospective procedure meant that prior to initiating a search for data (primary research studies), a specific set of criteria were established for

identifying studies to be included in the review and analysis.

In addition to including only studies that investigated the effects of a sign language intervention approach (sign language alone or total communication), studies included in this review also needed to meet several additional criteria, presented in Table 1.

Table 1. Criteria used for study selection

Design Criteria
Experimental or quasi-experimental group design, or experimental single subject design;
Statistical Criteria
a. Means, standard deviations, significance levels, proportions, or test of inferential analyses had to be reported for experimental and quasi-experimental group design studies; b. Single subject experimental design studies needed to use an ABA or multiple-baseline design and report a minimum of three data points for both the A and B conditions; c. All data had to be presented in a quantifiable form;
Participant Criteria
Participants were children with autism between 4 and 18 years of age;
Outcome Criteria
Sign or oral communication measured.

Study Retrieval Strategy

Studies for this review were collected by hand and electronic searches. An initial search was conducted using electronic databases (MEDLINE, CINHAL, ERIC, Exceptional Child Education Resources, Linguistics and Language Behavior Abstracts, PsycINFO) and books appropriate to the topic.

Reference lists of all studies retrieved from the electronic and hand searches were reviewed for any additional studies. The main source for the identified studies came from journals and papers presented at professional meetings.

Once a list of potentially appropriate studies was identified, full text copies were reviewed independently by the two authors for inclusion criteria; any differences in the judgment to include or exclude studies were resolved by discussion and consensus. Of 32 studies identified as possibly appropriate for inclusion, only eight studies met the criteria. These eight studies served as the basis for the summary and analysis reported in this work. A list of excluded studies is available from the authors.

Study Coding

After identifying the eight studies for inclusion, each was coded for research design, participant, and treatment characteristics. Coding of research design differentiated studies into two categories: group or single subject design. Coding of participant characteristics identified the number, age, gender, pre-treatment language ability, and referral setting of study participants. Coding of treatment characteristics identified treatment outcomes (i.e., sign or oral communication), number and length of sessions, type of treatment (sign only or total communication), fidelity of treatment implementation, and the effect type (training, generalization, or follow-up).

Statistic for Group Design Data

To assess the effectiveness of interventions that use group designs (e.g., experimental or quasi-experimental), the statistical metric of choice is typically an effect size metric referred to as a d statistic. The effect size d is calculated by subtracting the mean of the control group from the mean of the experimental group and then dividing by the average of the two groups' standard deviations for each outcome measured. The significance of the effect size is determined by noting whether or not the 95% confidence interval includes zero (0). If it does not (e.g., 95% CI= .14 to .68), the effect size d metric is said to be statistically significant. If the d value includes zero (e.g., 95% CI= -.14 to .68), the intervention effect is interpreted as not statistically significant, because it

means that the difference between the two groups' scores could be 0 (or lower). In general, a *d* value of less than .20 is interpreted as reflecting little impact of the treatment, a *d* value between .20 and .70 is interpreted as a moderate effect, and a *d* value of .70 or larger is interpreted as a strong intervention effect (Cohen, 1988).

**Statistic for Single Subject Design Data**

Single subject design outcome data require a different statistical metric to quantify the effects of intervention than found in group studies, namely the percentage of non-overlapping data (PND) (see Kazdin, 1982, & Mastropieri & Scruggs, 1985-86, and also arguments in Franklin, Allison, & Gorman, 1997; Scruggs & Mastropriori, 1998; Scruggs, Mastropriori, Forness, & Kavale, 1988). PND was used in this review to evaluate the size of the effects observed in single-subject investigations. The PND value is a representation of the proportion of non-overlapping data points across baseline and treatment measurements. It is calculated by counting the total number of data points measured during the treatment phase that exceed the most positive data point in the baseline phase. We use the term 'most positive data point' to reflect the possibility that a decrease in value may reflect a positive change (e.g., decrease in inappropriate behavior). For example, in Figure 1, the most positive data point on baseline is the fifth data point occurring on the 5th day of baseline. During treatment, seven of the 10 data points were greater than 4, thus the PND= 7/10 or 70%.

> **A PND of 90% or greater reflects a very effective treatment, whereas a PND between 70% and 90% is mildly effective and <70% is ineffective.**

The PND statistic is a straightforward and easily understood metric in which the higher the percentage, the more effective the treatment. Scruggs and Mastropieri (1998) have shown that when the PND reaches 90% or greater, the intervention is considered to be very effective. A moderate level of effect is produced when the PND is between 70% and 90%, whereas 50% to 70% is considered to be mildly effective or questionable. Any PND below 50% is said to be ineffective since the performance during intervention is no better than baseline performance at the level of statistical chance.

Figure 1: Sample data for PND calculation

PND=7/10=70%

SESSIONS
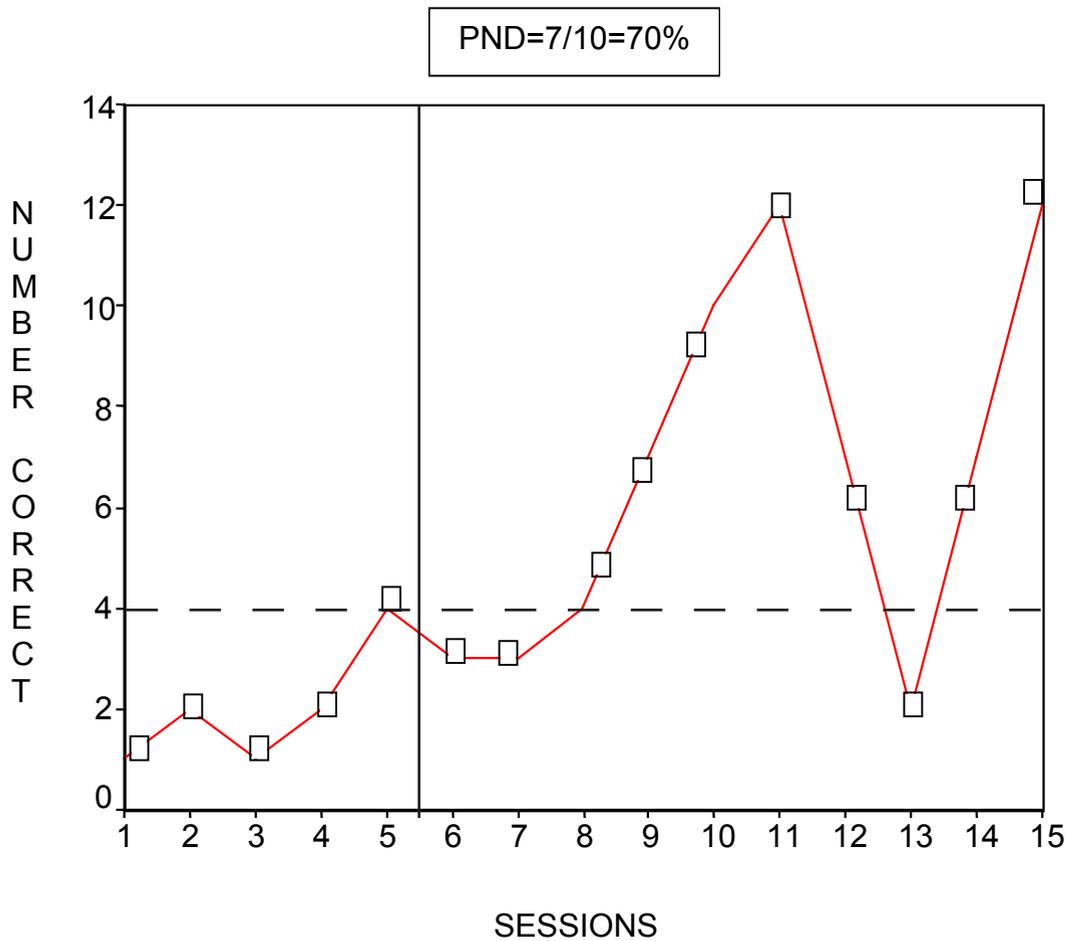
Results
Evaluating the Evidence

Data from the eight included studies were analyzed to assess the impact of sign or total communication intervention to improve the communication of participants on a pre- and post-intervention comparison. Of the eight studies meeting the inclusion criteria, one was an experimental group design study (Yoder & Layton, 1988), and the remaining seven studies utilized a single subject design.

Experimental Design Group Study

The only study employing a true experimental design (Yoder & Layton, 1988) compared the effects of four types of treatments: speech alone, sign alone, total communication, and alternating sign and speech. Sixty children with autism were randomly assigned to one of the four conditions, and each participated in 90 individual daily sessions of 40 minutes each. The outcome measure of interest was the number of spontaneous words the child used during the training sessions; echolalic or teacher prompted words were not included.

> **Yoder and Layton's (1988) study found the largest effect size contrast to reflect the difference between speech alone and sign alone, in favor of speech alone.**

Results of this study found that children used significantly more spontaneous words during the three treatment conditions including an oral component (speech alone, total communication, or alternating sign + speech) compared to the sign alone condition. Calculation of the effect size estimates (see Table 2) found that the largest effect occurred in the speech alone condition contrasted to the sign alone group. The d = .73 indicates that the group receiving the speech alone intervention had a mean performance that was nearly three-fourths of a standard deviation unit greater than the performance of the group receiving sign alone.

Table 2. Effect size (d), 95% confidence interval, and p value associated with the four treatment comparison conditions.

	d*	Upper	Lower	p value
Speech Alone vs Sign Alone	.73	- .04	1.50	.06
Speech Alone vs Alternating	.43	- .33	1.18	.25
Speech Alone vs Sign + Speech	.33	- .42	1.08	.37
Speech Alone vs Sign + Speech	.06	- .69	.80	.88
Sign Alone vs Sign + Speech	-.33	-1.08	.43	.38
Sign Alone vs Alternating	-.30	-1.05	.45	.42

*The d value reflects the effect of the first outcome to the second outcome (e.g., .73 is the advantage of the Speech Alone versus the Sign Alone condition)

In considering these findings, we must be careful in evaluating these statistics due to substantial variability in the performance of both groups, as shown by two additional measures. First, a significance test can be performed to determine the likelihood that the observed difference could have occurred by chance. For instance, the d = .73 is associated with a p value of .06, which falls short of the conventional acceptable value of <.05. Second, and perhaps more importantly, a confidence interval can be determined to estimate the upper and lower limits within which the true effect size is likely to reside. In this case, the d = .73 is associated with a 95% confidence interval of -.04 to 1.50. There are two critical features of this interval. First, it is very broad, showing a great deal of variability in the outcomes observed. Second, the lower boundary of the confidence interval is less than zero. This indicates that the probability that the true effect is greater than zero is less than 95%. Another way of saying this is that there is more than a 5% chance that the true effect for this comparison includes zero, meaning that neither sign alone nor speech alone resulted in a greater improvement of communication performance. Obviously, because all of the other observed effects shown in Table 2 are even smaller, none of these were statistically significant either. Yoder and Layton (1988) concluded that any advantages achieved by either the total communication or alternating sign + speech conditions might be due to either the verbal or the sign component of the training, and that it was not possible to quantify the independent effects of these two conditions.

In general, the data from this study offer little quantitative support for the use of sign alone or in conjunction with spoken language (total communication) to improve the spontaneous sign or oral communication of children with autism. The effect size analysis showed that none of the outcomes observed were statistically significant. Although the authors argued that their data support sign language training for children with autism, they did not in fact present commensurate analyses and statistical results to support their contentions.

Single Subject Design Studies
Design Characteristics

Seven single subject design studies were included in this review, and these were analyzed using PND analysis to quantify the effects of intervention. Four studies utilized a multiple baseline across behaviors design, two studies reported a multiple baseline across participants design, and one study applied a multiple baseline across settings design, as shown in Table 3.

Table 3. Participant and design characteristics

Study	Design	# of cases*	Mean Age	Setting	Pre Treatment Language Status
Carr et. al., 1978	MBL across Behaviors	M-4	13.8 yrs	Res	Vocalizations
Cohen, 1979	MBL across Behaviors	F-1	4.0 yrs	NR	Echolalic
Casey, 1978	MBL across Participants	M-3 F-1	NR	Ho	Echolalic
Schepis et. al., 1982	MBL across Setting	M-4	9.5 yrs	Res	Vocalizations/ Verbalizations
Carr & Kologinsky, 1983 Exp 1	MBL across Participants	M-3	11.0 yrs	Sch	Vocalizations/Signs
Remington & Clarke, 1983	MBL across Behaviors	M-1 F-1	12.5 yrs	Sch	M-Echolalic F-Vocalizations
Carr, et. al., 1987	MBL across Behaviors	M-4	14.0 yrs	Ho	Vocalizations

*M=male; F=female, Res=residential, Ho=home, Sch=school, NR=not reported, MBL=multiple baseline

A correlational analysis of the number of baseline and treatment data points with the PND for each study yielded a non-significant relationship between PND and baseline ($r = .055$, $p = .429$) and PND and treatment data points ($r = -.123$, $p = .345$). These findings indicate that no substantive relationship exists between the PND and baseline or treatment data points and the relationship does not introduce a statistically significant systematic bias in the overall analysis. [1]

Participant Characteristics

As seen in Table 3, a total of 22 children (F=3, M=19) ranging in age from 4 to 14 years of age (mean age=9;9) were treated in the seven analyzed studies. These children were referred from Home, School, Residential, and Unreported settings. The children's pre-treatment language ability ranged from limited vocalizations (n=9), to echolalic speech (n=6), to a mix of vocalizations/ verbalizations (n=4) or vocalizations with limited sign ability (n=3).

Treatment and Outcome Characteristics

A summary of treatment and study outcome characteristics is presented in Table 4. The following provides a description of the characteristics and the resultant effect of treatment.

Types of Interventions

For the seven studies, two different types of interventions were reported: sign only (n=3) and sign + speech (n=5). The total numbers for each type of treatment reflect the use of more than one treatment in one of the studies. The sign language systems used in training included American Sign Language (ASL, n=1; Ameslan, n=1), Signed English (n=1) and Not Reported (n=4).

Number and Length of Treatment Sessions

The overall length of the treatment program was reported in only two of the studies, at four weeks and seven weeks, respectively. The number of treatment sessions ranged from 3 to 10 sessions per week with the number of minutes per session ranging from 5 to 60 minutes. The number of overall treatment sessions per individual was highly variable, ranging from 3 to 72 sessions.

Table 4. Treatment and outcome characteristics

Study	type of treatment	form of sign language	measured outcome	# of baseline/data points	total # of sessions	# of sessions per week	time per session minutes	length of program	training PND
Carr et. al., 1978	sign + speech	ASL	Sign[1]	180/180	3-15	3-4	60	NR	100
Cohen, 1979	sign + speech	Ameslan	Prompted Sign + Speech[1] Unprompted Sign + Speech[1] Echolalic Speech Spontaneous Speech[1]	24/52 24/52 24/52 24/52	22 22 22 22	3 3 3 3	20 20 20 20	7 weeks 7 weeks 7 weeks 7 weeks	74 67 100 94
Casey, 1978	sign + speech	SE	Sign + Speech1	40/40	22		20	4 weeks	95
Schepis et. al., 1982	sign + speech	NR	Prompted Sign[1] Non-Prompted Sign[1]	90/160 90/120	22-72	5	NR	NR	74 29
Carr & Kologinsky, 1983 Exp 1	sign only	NR	Sign[1] Sign + Speech[1]	30/65	28	5	5-10	NR	100
Remington & Clarke, 1983	sign only sign + speech	NR	Sign[1] Sign + Speech[1]	30/30 30/30	NR	10*	15	NR	100 100
Carr et. al., 1987	sign only	NR	Sign[2]	76/143	4-23	3-5	45	NR	89

*2 sessions/day, [1]=single word, [2]=2 word combination, NR=Not Reported, ASL=American Sign Language, SE=Signed English, Ameslan=American Sign Language (Modified)

Treatment Fidelity

Treatment fidelity refers to the integrity of the treatment program; that is, was the treatment delivered as prescribed to the child? The possibility of a Type I error (a significant treatment effect but an unintended treatment variable was added to the treatment which may have contributed to the outcome) or Type II error (no treatment effect but the treatment was not implemented as intended) can increase when treatment fidelity is compromised. None of the included studies provided details of treatment fidelity; in addition, none of the studies included treatment descriptions in sufficient detail for replication by others.

Intervention Outcomes

In all studies, the effect of the intervention was measured for either oral, sign, or total communication outcomes. The average PND for treatment outcomes of the seven studies across all interventions was 80%, suggesting a moderate treatment effect. The average PND for oral communication was 60% (n=1), 87% for sign language only (n=5), and 84% for sign + speech (n=4) (total numbers reflect more than one outcome in several studies). On average, participants demonstrated a moderate degree of communication improvement regardless of the communication intervention strategy, although the largest effect (based on the PND) occurred for sign only. Only three of the seven studies presented generalization or follow-up data, and two of the three found that the participants were able to transfer the treatment condition learning to new environments. Due to the small number of studies presenting appropriate data, no further analyses were appropriate regarding generalization or follow-up data.

Discussion

The use of sign language to facilitate the communication of children with autism has been a topic of interest for many years. On the one hand, clinicians who teach sign language to children with autism may argue that the child at least is provided with a mechanism for communication, even if it may be limited in the breadth and sophistication of information. Further,

supporters might argue that the use of sign language serves as a mediator for the development of oral communication skills once the child recognizes the social value of language usage.

On the other hand, some clinicians may argue that the use of sign serves as a communicative crutch for which the child has little functional use, and that teaching children with autism to use sign may sidetrack the development of functional oral language skills required in the social and educational environment. Just as importantly, the detractors also argue that there is no substantive evidence that sign language improves communication beyond a one or two word sign structure, thus limiting the usefulness of sign language skills to a communication equivalent of labeling. Certainly, they would point out that the efficacy of sign language as a remedial program for children with autism is at best questionable.

The data from all sources presented in our review might be construed as offering a modicum of support for the use of sign language intervention. We identified seven single subject design studies that suggested an overall moderate treatment effect (PND=80%) for teaching communication skills to children with autism using signs only or a total communication strategy. However, because these were single subject studies and treatment fidelity was not documented, generalization of this conclusion beyond the study participants is unwarranted, even if the specific conclusion is correct. Additionally, none of the included studies provided sufficient information to replicate the intervention program in a clinical setting. Furthermore, evidence from the only group experimental design study (Yoder & Layton, 1988) showed no statistically significant effect size differences between any of the oral and sign combinations that they tested. Nonetheless, although these studies do not provide strong empirical support for the use of sign language intervention to promote communication in children with autism, it is also important to note that no evidence from the studies included in the meta-analysis suggested that using signs alone or in conjunction with speech was harmful or in any way contraindicated.

Several important limitations concerning the research examined

> **Although the effect of sign language intervention is moderate, a recommendation for clinical application is tempered by the quality of research from which the data are drawn.**

warrant note. First, no study included measures of intervention fidelity. Second, few studies included measures of generalization. Third, few studies provided adequate detail of the intervention, thus making replication difficult if not possible. Fourth, only one group design experimental study met inclusion criteria and this type of design is considered among the strongest in terms of establishing causality. For the one group experimental study available, the effect size estimates showed there to be no statistically significant differences between any of the four conditions. Thus, while eight studies were available for this review, the conclusions drawn from our analyses must be tempered by the quality of the research available at this time.

This review spotlights the glaring shortage of high quality research needed to inform any discussion of the merits of teaching sign language to children with autism. In the intervening 18 years since Yoder and Layton (1988) called for experimental research on this topic, little appears to have been accomplished. Thus, clinicians' use of a sign language approach to enhance the communicative competence of children with autism must be considered in light of (a) the absence of conclusive group experimental design evidence to corroborate the single subject design findings, and (b) the absence of a discussion of intervention fidelity in all studies reviewed. From a programmatic and policy implementation point of view, the single subject research offers limited support for the use of sign language for children with autism. Considering the overall quality of the available research we would suggest that there are insufficient data to advocate for the use of sign language either alone on in combination with oral language as a method for substantially improving communication in children with autism.

Evidence-Based Practice Recommendation

Evidence-based practice requires clinicians to integrate the scientific, objective, and quantifiable data available in the research literature into the clinical decision-making process. Evidence-based practice should allow professionals to consider a variety of sources of information in light of the client's needs and the situation of the individual being treated. In this review, the evidence on the use of sign language with children with autism provides limited support for its

concentrated application for children with autism, as there is little compelling evidence that sign language provides substantial improvements in either oral or sign language communication. The modest effects reported by single subject studies coupled with the absence of even a few well controlled group studies only serves to suggest that either (a) the research community views this area of intervention as having limited usefulness, or (b) the clinical community has not found sign language to be of a substantial a value so as to press for more and better research. This review indicates that there is a need for high quality primary research that will provide the scientific basis for the effective clinical application of sign language intervention for children with autism.

[1] One of the contentious issues in systematic review methodology centers on the presence of bias at several different points in the process. It could be argued that a potential bias might have occurred as a result of substantial differences in the number of baseline and data points reported, such that fewer baseline points might result in a greater effect and thus a larger PND.

References

Carr, E.G., Binkoff, J.A., Kologinsky, E., & Eddy, M. (1978). Acquisition of sign language by autistic children. I: Expressive labeling. *Journal of Applied Behavioral Analysis, 11,* 489-501.*

Carr, E.G., Kologinsky, E., & Leff-Simon, S. (1987). Acquisition of sign language by autistic children. III: Generalized descriptive phrases. *Journal of Autism and Developmental Disorders, 17, (2)* 217-229.*

Carr, E.G., & Kologinsky, E. (1983). Acquisition of sign language by autistic children II: Sponaneity and generalization effects. *Journal of Applied Behavior Analysis, 16,* 297-314.*

Casey, L.O. (1978). Development of communicative behavior in autistic children: A parent program using manual signs. *Journal of Autism and Childhood Schizophrenia, 8, (1)* 45-59.*

Cohen, M. (April, 1979). The development of language behavior in an autistic child using a total communication approach. *A paper presented at the Annual International Convention,* The Council for Exceptional

Children, Dallas, TX.*

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd Ed.)* Hillsdale, NJ: Erlbaum.

Creekmore, N.N., (1982). Use of sign alone and sign plus speech in language training of nonverbal autistic children. *Journal of the Association for the Severely Handicapped, 6,* 45-55.

Franklin, R.D., Allison, D.B., & Gorman, B.S. (1997) *Design and analysis of single-case research.* Mahwah, NJ: Erlbaum.

Goldstein, H. (2002). Communication intervention for children with autism: A review of treatment efficacy. *Journal of Autism and Developmental Disorders, 32,(5)* 373-396.

Kavale, K.A., Mathur, S.R., Forness, S.R., Quinn, M.M., & Rutherford, R.B. (2000). Right reason in the integration of group and single subject research in behavioral disorders. *Behavioral Disorders, 25, (2)* 142-157.

Kazdin, A.E. (1982). *Single-case research designs: Methods for clinical and applied settings.* New York: Oxford University Press.

Kiernan, C. (1983). The use of nonvocal communication techniques with autistic individuals. *Journal of Child Psychology and Psychiatry, 24, (3)* 339-375.

Miranda, P. (2002). Autism, augmentative communication, and assistive technology: What do we really know? *Focus on Autism and Other Developmental Disabilities, 16, (3)* 141-151.

Mastropieri, M.A., & Scruggs, T.E. (1985-1986). Early intervention for

socially withdrawn children. *Journal of Special Education, 19,* 429-441.

Remington, B., & Clarke, S. (1983). Acquisition of expressive signing by autistic children: An evaluation of the relative effects of simultaneous communication and sign-alone training. *Journal of Applied Behavior Analysis, 16,* 315-328.*

Sackett, D.L., Straus, S.E., Richardson, W.S., Rosenberg, W., & Haynes, R.B. (2004). *Evidence-based medicine: How to practice and teach EBM.* New York, NY: Churchhill Livingston.

Schepis, M.M., Reid, D.H., Fitzgerald, J.R., Faw, G.D., van den Pol, R.A., & Welty, P.A. (1982). A program for increasing manual signing by autistic and profoundly retarded youth within the daily environment. *Journal of Applied Behavior Analysis, 15,* 363-379.*

Scruggs, T.E., Mastropieri, M.A., Forness, S.R., & Kavale, K.A. (1988). Early language intervention: A quantitative synthesis of single-subject research. *Journal of Special Education, 22,* 259-283.

Scruggs, T.E., & Mastropieri, M.A. (1998). Summarizing single-subject research. *Behavior Modification, 22,* 221-143.

Tawny, J.W., & Gast, D.L. (1984). Single subject research in special education. Columbus, OH: Merrill Publishing Company.

Yoder, P.J. & Layton, T.L. (1988). Speech following sign language training in autistic children with minimal verbal language. *Journal of Autism and Developmental Disorders, 18, (2)* 217-229.*

*Studies meeting inclusion criteria for analysis

Improving Communication for Children with Autism:
Does Sign Language Work?

Jamie B. Schwartz, PhD
Chad Nye, PhD
University of Central Florida