University of Alberta

EXAMINATION OF THE METHODOLOGICAL QUALITY AND EFFECTIVENESS OF SINGLE-CASE STUDIES ON SOCIAL SKILLS INTERVENTIONS OF CHILDREN WITH AUTISM SPECTRUM DISORDERS

by

SHIN-YI WANG

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

SPECIAL EDUCATION

DEPARTMENT OF EDUCATIONAL PSYCHOLOGY

SPRING 2013
Edmonton, Alberta

Abstract

This dissertation includes three studies investigating the methodological quality and treatment effectiveness of single-case research on social skills interventions for children with autism spectrum disorders. Study One developed a list of quality indicators that can be used to examine methodological quality of studies that use single-case research methods. Study One presents the rationale for each indicator and demonstrates how these indicators can be applied. Study Two examined the effectiveness of two different intervention approaches, peer-mediated and video-modeling approaches, by using hierarchical linear modeling. Both intervention approaches were found to significantly improve social behavior of participants with autism spectrum disorders, the two approaches did not differ in their effectiveness, and their effectiveness was moderated by the age of the participants. Study Three investigated treatment effects of 115 single-case studies on social skills interventions for children with autism spectrum disorders and examined the impact of three moderators – age, length of intervention, and research design – on the treatment effectiveness. An overall large mean effect size was found. Only one of the examined moderators, type of research design, moderated the intervention effects: the studies that adopted multiple baseline or reversal design tended to have better outcomes than the studies that adopted other types of single-case research designs. Together, the three studies increase our understanding of the research quality and intervention effectiveness of single-case studies on social skills interventions for individuals with autism spectrum disorders and help to promote evidence-based practices in this field.

ACKNOWLEDGMENTS

The idea for this three-paper dissertation came from a course taught by Dr. Rauno Parrila that I took in the first semester of my PhD program. The process of writing the papers was much more demanding that I could image, and I am very much indebted to my wonderful committee members for the guidance and support provided to me during the process.

Dr. Parrila has been especially helpful for my research and PhD program. He has been very patient and understanding, and provided me with all the supervision that I need to become a researcher. Dr. Ying Cui has been very caring and encouraging and helped me to explore new methods for the research. By being there for discussions with me, Dr. Christina Rinaldi helped me through a difficult time. The learning experience with Dr. Dick Sobsey facilitated my progress of learning to think as a researcher. Dr. Joyce Magill-Evans, the wonderful mentor at University of Alberta autism research group, provided me with guidance and support. All of my committee members represent different great scholar models for me, and remind me of an old Chinese saying: “A teacher is to disclose the fundamental principles between oneself and the others, to transmit the knowledge and skills for our careers, and to motivate us to think and resolve puzzles in the learning process.”

Finally, I am grateful to my family for the support that they have provided throughout my academic journey. With their company and support, all the hardships were winnable.

TABLE OF CONTENTS

ACKNOWLEDGMENTS ii

LIST OF TABLES vii

LIST OF FIGURES viii

ABSTRACT ix

CHAPTER I: INTRODUCTION.....1

Autism Spectrum Disorders and Social Interaction 2

Social Skills Interventions for Individual with Autism Spectrum Disorders.....7

Synthesis Methods for Single-Case Research on Social Skills Interventions for Individuals with Autism Spectrum Disorders 13

Quality criteria for single-case research.....14

Meta-analysis of single-case research.....15

Current Dissertation 23

References.....26

CHAPTER II: QUALITY INDICATORS FOR SINGLE-CASE RESEARCH ON SOCIAL SKILL INTERVENTIONS FOR CHILDREN WITH AUTISTIC SPECTRUM DISORDER.....37

Introduction.....37

Quality Indicators.....39

Primary quality indicators.....40

Participants.....40

Settings and materials used for social skill

training.....43

Independent variables.....43

Dependent variables.....44

Research designs.....45

Secondary quality indicators.....46

External validity.....46

Social validity.....47

Examples of How the Quality Indicator Checklist Can Be Used….....48

Target papers….....48

Scoring of individual papers….....50

Examination of the results by quality indicators.....51

Discussion.....52

References.....61

CHAPTER III: EXAMINING THE EFFECTIVENESS OF PEER-MEDIATED AND VIDEO-MODELING SOCIAL SKILLS INTERVENTIONS FOR CHILDREN WITH AUTISM SPECTRUM DISRODER: A META-ANALYSIS IN SINGLE-CASE RESEARCH USING HLM.....65

Introduction.....65

Method….....71

Identification and selection of reviewed papers….....71

Coding the studies….....73

Transformation of the data points for HLM and effect

sizes.....74

Examining effect sizes with HLM.....74

Results.....76

Discussion.....77

References.....86

CHAPTER IV: META-ANALYSIS OF SOCIAL SKILLS INTERVENTIONS OF SINGLE-CASE RESEARCH FOR INDIVIDUALS WITH AUTISM SPECTRUM DISORDER: RESULTS FROM THREE-LEVEL HLM.....93

Introduction.....93

Method.....99

Identification and selection of reviewed papers.....99

Quality coding.....100

Effect size calculations.....101

Transformation of the data points.....102

Statistical analyses.....102

Results.....106

Quality indicators.....106

Effect sizes.....106

Discussion.....108

References.....124

CHAPTER V: GENERAL DISCUSSION.....147

Limitations.....156

Conclusions.....159

References.....161

Appendix A: The Quality Indicator Checklist for Single-Case Research in ASD.....165

LIST OF TABLES

Table 2.1. The Initial Checklist of Quality Indicators for Single-Case Studies of Social Skills Training for Children with ASD.....56

Table 2.2. Summary of the Reviewed Papers.....57

Table 2.3. The Percentage of Papers Meeting the Criteria of Each Primary and Secondary Quality Indicators.....60

Table 3.1. Summary of the Reviewed Papers.....83

Table 3.2. Results of 3-Level HLM Final Estimation of Fixed Effects (with Robust Standard Errors) 84

Table 4.1. Summary of the Reviewed Papers.....115

Table 4.2. Quality Indicators for Quality Examination.....119

Table 4.3. Results from 3-Level Fully Unconditional HLM Model (With Robust Standard Errors) 120

Table 4.4. Results from Conditional 3-Level HLM Model Including Intervention Length as Level1 Predictor (With Robust Standard Errors).....121

Table 4.5. Results from Conditional 3-Level HLM Model Including Gender and Age Length as Level 2 Predictor (With Robust Standard Errors) 122

Table 4.6. Results from Conditional 3-Level HLM Model Including Quality Indicators as Level 3 Predictor (With Robust Standard Errors) 123

LIST OF FIGURES

Figure 3.1. The interaction effect of treatment type and age on effect sizes.....85

CHAPTER I

INTRODUCTION

The impairment of social interaction is recognized as one of the core deficits for children with autism spectrum disorder (ASD) (Carter, Davis, Klin, & Volkmar, 2005; Krasny, Williams, Provencal, & Ozonoff, 2003; Volkmar & Klin, 2005; White, Keonig, & Scahill, 2007). The demand for social skills interventions (SSIs) for children with ASD has been increasing not only because of the predominance of social impairments in ASD but also due to the escalating prevalence of ASD and increasing opportunities for social interaction in inclusive education settings. A wide variety of different models have been proposed and adopted in studies of SSIs for children with ASD. Several studies (e.g., Koegel, Werner, Vismara, & Koegel, 2005; Laushey & Heflin, 2000; Liber, Frea, & Symon, 2008) have claimed that their interventions resulted in significant progress in social behavior in their participants with ASD, whereas others (e.g., Barnhill, Cook, Tebbenkamp, & Myles, 2002；Marriage, Gordon, & Brandt, 1995) have demonstrated more limited or inconsistent benefits for their participants. How to systematically examine the methodological quality and treatment effectiveness of these studies has become a critical issue for researchers, practitioners, and families who are interested in this field as systematic examination of these studies can inform researchers, practitioners, and families how well these studies have been carried out and what the overall outcomes of these SSIs are. With this information, researchers may modify and improve SSIs, practitioners can provide more effective SSIs to individuals with ASD, and

families can make better choices of SSIs that can benefit their children with ASD. However, there are limited studies that have examined how well the SSI studies have been implemented and what their overall treatment effects are.

This dissertation starts with an introduction that presents a literature review of autism spectrum disorders and social interaction, social skills interventions for individuals with autism spectrum disorders, and synthesis methods for single-case research on social skills interventions for individuals with autism spectrum disorders. Chapters 2, 3, and 4 present the three papers. The dissertation concludes with a General Discussion (Chapter 5).

Autism Spectrum Disorders and Social Interaction

Given the differences across children with autism, the concept of a spectrum was introduced (Hall, 2009, chap. 1) in order to highlight the variability. The term autism spectrum disorders (ASD) has been used unofficially to represent autism and autistic related disorders listed in the International Classification of Diseases, Tenth Revision (ICD-10) (World Health Organization [WHO], 1992). A synonymous term, pervasive developmental disorders, is used in the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR) (American Psychiatric Association [APA], 2000). Before the forth-coming publication of DSM, Fifth Edition (DSM-V) (APA, 2012) in May 2013, ASD consists of a group of developmental disorders including autistic disorder or childhood autism, pervasive developmental disorders not otherwise specified (PDD-NOS) or atypical autism, Asperger disorder or syndrome, Rett disorder or syndrome, and childhood disintegrative disorder or other childhood

disintegrative disorder (Gillberg, 2006; National Institute of Mental Health, 2009; Spector & Volkmar, 2006). However, in DSM-V (APA, 2012), Rett disorder or syndrome will no longer be included in ASD category, and the terms autistic disorder, PDD-NOS, Asperger's disorder, and childhood disintegrative disorder will no longer be used but integrated to a single ASD category. According to the criteria in DSM, Fourth Edition (DSM-IV) (APA, 1994) and ICD-10 (WHO, 1992), individuals with ASD are characterized by significant impairments in social interaction and communication and the presence of restricted or repetitive patterns of behavior and interests. The domains of social interaction and communication are combined in DSM-V (APA, 2012) and at least five of seven criteria in DSM-V, instead of six out of twelve listed in DSM-IV-TR, will be required in order to meet the diagnostic requirement of ASD (McPartland, Reichow, & Volkmar, 2012).

The estimate of prevalence of children with ASD has dramatically increased from 4 – 5 in 10000 to 6 in 1000 children during the past few decades (Curran & Bolton, 2009; Hall, 2009; Wolff, 2004). The prevalence rate of ASD can vary depending on the conceptualization and diagnostic criteria used (Baird et al., 2006; Chakrabarti & Fombonne, 2005; Charman, 2002; Volkmar, Lord, Bailey, Schultz, & Klin, 2004). Center for Disease Control and Prevention (CDC, 2012) estimated ASD prevalence as 1 in 88 children based on the data collected during 2008 through the Autism and Developmental Disabilities Monitoring (ADDM) Network. In contrast, the CDC estimate was 1 in 150 children based on the ADDM data collected in 2000. Curran and Bolton (2009), Fombonne (2005), and

Smith (2008) indicated that the increasing rate was most likely due to the broader definition and more comprehensive diagnosis of ASD. As well, the awareness of professionals and the availability of supportive resources have also contributed to the increase in ASD prevalence (Hall, 2009).

Although precise biological mechanisms have not been found for ASD, recent studies support the significance of neurobiological mechanisms in the etiology of ASD (Barton & Volkmar, 1998). The association between ASD and certain medical problems and the higher prevalence rates in siblings and family members support the significance of genetic influence (Curran & Bolton, 2009). Currently, studies in pathogenesis emphasize the important roles of brain development and the interaction between genetic and environmental risk factors (Dawson, 2008; Hall, 2009; Volkmar & Klin, 2005). Dawson hypothesized that altered behavioral patterns in individuals with ASD, resulting from genetic and environmental risk factors at early sensitive stages, interfere with their brain development and keep them from engaging with others.

Difficulties in social interaction have been recognized as a core problem in individuals with ASD since the initial cases reported by Kanner (Carter et al., 2005; Wing, 1991). Problems with social interaction rather than language and behavioral difficulties are more pervasive in children with ASD and differentiate them from children with other diagnoses (Carter et al.; Krasny et al., 2003). National Research Council (NRC) (2001) indicated that children with ASD often have social problems in the areas of imitation, social initiations and responses, play, and communicative behavior. Typical children can naturally scaffold their

learning through observing and imitating adults and peers. However, imitating and learning from others is very limited in children with ASD. Compared with children at a similar age, children with ASD are either unable, indifferent, or inept when it comes to imitating others' behavior, interacting with others both verbally or nonverbally, joining peers' play, and adopting verbal or body language for communication facilitation (NRC, 2001). As well, social reciprocity with adults and peers and social orientation to faces and eye gaze are difficult tasks for people with ASD (NRC, 2001). Cotugno (2009) also pointed out that the social difficulties of people with ASD can be recognized in their lack of ability to understand others' intentions and social rules, respond to social contact (e.g., verbal or nonverbal communication) appropriately, and share feelings and interests with others. Social interaction is a complex process that requires comprehending others' actions and taking corresponding actions simultaneously. Because of the difficulties in initiating contacts as well as attending to and reading social cues such as facial expression, making appropriate responses and maintaining to-and-fro interaction are challenging for people with ASD.

It is readily understandable why the establishment and maintenance of peer relationships is difficult for children with ASD. Even when their parents and inclusive education settings provide sufficient opportunities for interaction with other children, children with ASD are less likely to establish successful peer relationships due to the lack of social skills and understanding (Kasari & Rotheram-Fuller, 2007). While peers are around, children with ASD make fewer efforts to engage with and to respond to them. Because individuals with ASD

frequently lack distinct understanding of the differences between self and other, the abilities to take others' perspectives and integrate the information of social context for social interaction, and the methods to attune self with others (Krasny et al., 2003), they are more likely to fail to interact effectively and form joyful reciprocal affiliations with peers. Moreover, social interaction becomes a more serious problem for individuals with ASD as they age (Kasari & Rotheram-Fuller, 2007; White, Keonig, et al., 2007). As children with ASD move up from one grade to another, they are more likely to be excluded from social networks because there are more complex social rules and peer relationships for older children, who demand more reciprocal interaction and prefer to interact with typical peers. Peer rejection and the sense of isolation can be linked with poor self-esteem and further issues in social adjustment (Krasny et al., 2003).

Varied psychological theories, including theory of mind, weak central coherence, and executive dysfunction have attempted to explain the underlying psychological process for social deficits in ASD children (Cotugno, 2009; Konstantareas, 2006; Rajendran & Mitchell, 2007; Reichow & Volkmar, 2010; Spector & Volkmar, 2006). Theory of mind hypothesizes that individuals with ASD fail to recognize that others' minds are different from theirs and have difficulties viewing things from others' perspectives (Baron-Cohen, Leslie, & Frith, 1985), or that their development of the abilities to read others' minds are delayed which is recognized as mindblindness (Baron-Cohen, 2008). Weak central coherence assumes that people with ASD are weak in obtaining the wholeness of the meaning or picture and tend to focus on the details as they

process information (Happé & Frith, 2006). Executive dysfunction was proposed because some core symptoms in people with ASD are similar to problems in patients with brain injury who have difficulties in executive function such as planning and flexibility (Hill, 2004). Although these theories can explain some difficulties in people with ASD, there remain limitations especially in terms of lack of a developmental perspective (Rajendran & Mitchell, 2007).

In sum, ASD currently represents several categories of childhood-onset disorders and syndromes with the major difficulties in social interaction, including the establishment of reciprocal affiliations and interpersonal relationships. Biological and environmental factors may both play significant roles in the etiology of ASD. Theory of mind, weak central coherence, and executive dysfunction are the major psychological theories that attempt to explain the social deficits in children with ASD.

Social Skills Interventions for Individuals with ASD

Since social deficits are recognized as one of the core impairments for individuals with ASD (Carter et al. 2005; Krasny et al., 2003; Volkmar & Klin, 2005; White, Keonig, et al., 2007), great efforts are devoted to developing social skills interventions (SSIs). Furthermore, the demand for SSIs is increasing due to the extension and progress of the diagnostic criteria that have helped identify more cases with the core deficit in social interaction such as high function autism (HFA), PDD-NOS or atypical autism, and Asperger disorder. Additionally, urgent need for SSIs arise as movement towards more inclusive education settings has increased the chances for students with ASD to socialize with peers and the social

difficulties of individuals with ASD become more apparent while they make more frequent contacts with others.

Because of heterogeneity of ASD, the difficulties of social interaction in individuals with ASD can vary widely. Therefore, SSIs for individuals with ASD may address different target behaviors ranging from simple component skills to multifaceted behaviors that require integration of several skills. Depending on the functional level, developmental stage, and the needs of individuals with ASD, the SSI target behaviors may include eye contact, verbal or nonverbal initiation, verbal or nonverbal response, greeting, waiting in group, pretend play, collaborative play, turn-taking, joint attention, friendship-building, social problem-solving, and understanding others' perspective.

Although greater attention has been given lately to SSIs for individuals with ASD (Reichow & Volkmar, 2010), social skills interventions have been seen as less promising compared with interventions for cognition, communication, or behavioral problems (NRC, 2001; White, Keonig et al., 2007). SSIs can be very challenging because individuals with ASD tend to have multiple difficulties in cognition, communication, social interaction, and behaviors (Krasny et al., 2003). Moreover, individuals with ASD benefit minimally from general SSI strategies and exposure to peers. For instance, many children with ASD do not actively observe and imitate peers' behaviors and they often have difficulties in applying the learned skills to different settings. Therefore, Krasny et al. (2003) suggested SSIs require taking into account the multiple deficits and personal characteristics of the participants as well as generalization issues.

Given the heterogeneous characteristics of ASD and uncertainties about its etiology, it is reasonable to expect the increasing varieties of intervention models in recent studies. Additionally, the growing number of individuals with ASD as well as the escalating demands for funding for ASD treatments has encouraged various researchers and professionals to put their efforts into the development of training programs. Depending on the characteristics of the SSIs, they may be categorized into behavioral orientation (i.e, reinforcement, prompts), peer-mediated training, social story, pivotal response training, joint attention training, and buddy system (Bass & Mulick, 2007; Matson, Matson et al., 2007; Scattone, 2007).

SSI models can also be categorized based on who is directing the intervention (e.g., adult, child with ASD, or peer) or what the theoretical orientation is (e.g., behavioral, neobehavioral, or developmental approaches). For instance, social skills interventions can be divided into three different strategies – adult-directed, child-directed, and peer-mediated – depending on who is responsible for leading the intervention procedure (NRC, 2001; Spector & Volkmar, 2006). In adult-directed programs, adults such as therapists, teachers, or parents, implement the training to children with ASD. In child-directed programs, children with ASD develop their social skills as adults respond to and build children's social potential upon the children's spontaneous behavior. In peer-mediated programs, typically developing children, used as scaffolded supports, are trained to prompt and elicit appropriate social behavior from children with ASD. Once the training is completed, the peers are asked to apply

those strategies while interacting with children with ASD in order to promote their social interaction. The peer-mediated approach tends to be the most empirically supported model of social skills intervention for children with ASD (Apple, Billingsley, & Schwartz, 2005; Bass & Mulick, 2007). Its treatment effectiveness was established in increasing communication, interpersonal, and play behaviors of individuals with ASD in the 2009 report of the National Standard Project by National Autism Center (National Autism Center, 2009). Training peers to deliver the intervention emerged in 1960 as people recognized the limitations in adult-mediated training (Strain, Schwartz, & Bovey, 2008), and this approach benefited young children with ASD and ASD children with minimal ability (NRC, 2001). Strain et al. (2008) also suggested that peer-mediated intervention is viewed as a more natural and generalizable way of learning than the intervention instructed directly by adults. Furthermore, Kasari and Rotheram-Fuller (2007) argued that some programs with a combination of different strategies such as adult-directed and peer-mediated interventions had successfully improved social performance of children with ASD.

Social skills interventions can also be embedded in comprehensive programs that are guided by varied theoretical orientations such as behavioral, neobehavioral, or developmental approaches (NRC, 2001). Behavioral programs, known as applied behavioral analysis (ABA), implement interventions based on the principles of behaviorism. Behavioral strategies such as modeling, reinforcement, imitation, and prompts have been widely applied to varied types of intervention programs (Reichow & Volkmar, 2010). Neobehavioral programs such

as the Walden program, the Learning Experiences-Alternative Program, and the pivotal response training modify the implementation of behavioral intervention by embedding the intervention in the natural setting or incidence of social interaction (NRC, 2001). Developmental approaches such as Greenspan and Wieder's intervention program, known as Floortime or DIR (Developmental, Individual-difference, Relationship-based) treatment (NRC, 2001), have been viewed as being on the other end of the continuity of intervention philosophy. Development-oriented programs stress the importance of following child's lead and focus on the development of interpersonal relationship of children with ASD.

In addition to the models mentioned above, there are other SSI methods for individuals with ASD. Social stories were developed by Carol Gray in 1991 and are used to facilitate the social performance of children with ASD by writing down and illustrating social rules. Although positive gains were reported in a few studies, a more comprehensive review is still required to support the effectiveness of social stories. Video-modeling intervention, another SSI approach, has also been recognized as an effective method to improve the social-communication skills in children with autism (Apple et al., 2005; Bellini & Akullian, 2007), and autistic disorder, PDD-NOS, and Asperger syndrome (National Autism Center, 2009). The utilization of video media can provoke children's interest and be a natural way for children to learn skills. In this approach, children with ASD are asked to observe a video of proper social interaction and to mimic the social behavior demonstrated in the video. Additionally, SSI is delivered in comprehensive programs such as the Early Start Denver Model, and Treatment

and Education of Autistic and Communication-Handicapped Children (TEACCH) with coherent strategies matching the corresponding model. As well, there are a variety of social training programs that combine varied strategies listed above. However, while some social skills interventions are proving to be promising, other intervention approaches, such as holding therapy and animal therapy, are not supported by empirical evidence (Simpson, 2005). According to *Autism Spectrum Disorders: Guide to Evidence-based Interventions* (2012) that Missouri Autism Guidelines Initiative published after reviewing six systematic reviews of ASD treatment studies, some comprehensive programs and specific intervention methods had evidence to support their effectiveness in improving social behavior of individuals with ASD, although they were not necessarily supported by all of the reviews. These included comprehensive behavioral intervention programs, naturalistic interventions, peer-mediated interventions, and social narratives. In contrast, complementary and alternative medicine treatments were found to lack sufficient evidence of social benefits.

In sum, increasing varieties of SSIs have been developed for individuals with ASD, possibly due to the heterogeneous symptoms of ASD and the increasing number of individuals with ASD. However, not all SSIs have shown promise in improving the social skills of children with ASD. More comparative and systematic reviews of the varied SSIs are in demands as these reviews can help families, researchers, and other people in the field get a clearer picture of research quality and treatment outcomes of the SSIs.

Synthesis Methods for Single-Case Research on Social Skills Interventions for Individuals with Autism Spectrum Disorders

With the increasing number of individuals diagnosed with ASD and varieties of SSI models being used to treat with social skills deficits, the quality and efficacy of these models has become a great concern (Cotugno, 2009; Krasny et al., 2003). The examination of the varied models is critical because it can provide important information to desperate parents in order to prevent their waste of energy and money on less promising interventions (Gillberg, 2006) and the examination can come up with scientifically-supported information on interventions for professionals, families, and resource allocators. Moreover, the examination of SSIs provides opportunities to investigate the factors that can lead to the best fit between the child and the intervention, and thus allow the child to benefit maximally from the intervention and the allotment of treatment funding to be cost-effective.

Matson, Matson et al. (2007) reviewed social skills intervention studies for children with autism and indicated that 90% of the intervention studies employed single-case research methods. Similarly, over 80 % of the SSIs studies reviewed recently by Reichow and Volkmar (2010) adopted single-case research methods. Single-case research is one form of quasi-experimental research that measures the behavior of one or more participants repeatedly over time. Single-case research has been widely applied to studies in the fields of clinical psychology and special education for several reasons. First of all, the researchers and clinicians in the fields of clinical psychology and special education tend to work with participants

who are relatively few or of relatively low prevalence, thus making random assignment of participants to groups difficult. In addition, assigning clients or students to the control group may violate their needs for early intervention and raise ethical concerns. More practical reasons why clinicians and educators are more likely to be attracted to single-case research designs is the relatively low cost of single-case studies and the resulting in-depth information of participant responses (Van den Noortgate & Onghena, 2003a, 2007).

Because most of SSI studies for individuals with ASD have adopted single-case research methods, the syntheses of the studies using this approach must consider what kinds of synthesis tools and methods can be used to appropriately evaluate the methodological quality and the treatment effects of the single-case interventions studied. Generally speaking, study synthesis begins with the examination of methodological quality. For example, Horner, Carr, Halle, McGee, Odom, and Wolery (2005) proposed a list of quality criteria that can be used to evaluate multiple dimensions of single-case research including internal validity, external validity, and social validity.

Quality criteria for single-case research.

Internal validity refers to the extent to which studies can manage the confounding factors and demonstrate causal relationship between an intervention and the observed outcomes. The quality indices used to examine internal validity of single-case research include items such as does the study provide sufficient information about the participants, settings/material, and the independent and dependent variables; were the variables manipulated and measured faithfully; did

the study collect sufficient data points; does the data represent a stable pattern, and so on. External validity is related to how well the intervention results can be generalized across various time frames, settings, or participants. Therefore, the external validity of the study can be established if the study earns credit on the quality indicator of generalizability. The more that the intervention outcomes can been generalized to other participants or to different settings or time, the better the external validity of the study will be. Social validity denotes how well the intervention outcomes and their importance have been recognized. Quality indicators such as positive reports collected from significant others of the participants can be used to assess social validity of the study.

Lord et al. (2005), Smith et al. (2007), and Reichow, Volkmar, and Cicchetti, (2008) claimed that there are specific criteria for the studies on SSIs for individuals with ASD. Therefore, adding other quality indicators may be required for the evaluation of these studies. Additional quality indicators, such as use of standardized instruments for diagnosis, information on the peers and interventionists, the criteria for the percentage of the sessions used to examine the inter-rater agreement, utilization of multiple baseline or reversal design, the number of data points in the baseline and intervention phases, and the use of blind agents for establishing social validity were stressed because of their significant roles in research quality of single-case research on SSIs for individuals with ASD.

Meta-analysis of single-case studies.

In addition to the investigation of methodological quality, the synthesis of single-case SSIs for individuals with ASD also attaches importance to the

examination of intervention effectiveness. The evaluation of effectiveness of studies using single-case research typically relies on visual analysis or examination of the outcome graphs (Brossart, Parker, Olson, & Mahadevan, 2006; Miller, 2005; Morgan & Morgan; 2009; Parker & Hagan-Burke, 2007; Van den Noortgate & Onghena, 2003a). Visual analysis of the graphs means visually inspect the variability, trend, and overlap of the data points between or within baseline and intervention phases (Miller, 2005; Morgan & Morgan; 2009). Proponents of visual analysis support visual analysis for its in-depth information, convenience, low error rates, and conservatism in detecting effective treatments. However, visual analysis is more suitable for detecting changes in the measured behavior when the intervention effect is large and there is a stable baseline (Crosbie, 1999). Concerns arise for its subjectivity and applicability in comparison of multiple studies. Additional concerns have focused on the reliability and validity of visual analysis (Beeson & Robey, 2006; Brossart et al., 2006; Morgan & Morgan; 2009). Visual analysis can lead to misinterpretation of the efficacy of the implemented program when there are issues such as few data points, rater biases, or slight behavioral changes (Newell & Burnard, 2006). Furthermore, applying visual analysis to the comparison of treatment effects of multiple studies by asking different raters to read numerous graphs at one time is a complex task that is susceptible to biases. Inter-rater agreement between experts and well-trained raters has ranged from moderate to low and further calls into question the validity of visual analysis (Brossart et al., 2006). As a result, meta-analysis, a quantitative synthesis using statistical methods, has been

suggested as it can come up with more objective and comprehensive outcomes for the synthesis of multiple studies (Van den Noortgate & Onghena, 2003a).

Meta-analysis is recognized as a quantitative and objective tool for assessing the magnitude of intervention effects through the synthesis of a set of studies. According to Glass (1976, p. 3), meta-analysis uses diverse statistical methods to synthesize the quantitative outcomes of many independent studies. Effect sizes, an index of treatment success that reflects the magnitude of the effect of independent variable(s) on dependent variable(s), are most often used in meta-analytic studies and have been viewed with increasing credibility, especially among research professionals, as precise, accountable, and objective. As well, effect sizes, free of dependence upon sample size and measured on a continuous scale, can be used for the systematic comparison and summary of study outcomes. Meta-analysis can support visual analysis in detecting effect sizes especially when there is no stable baseline or the behavioral change is minimal, and also can help identify moderators of treatment effectiveness (Brossart et al., 2006; Van den Noortgate & Onghena, 2003a). However, meta-analysis is traditionally applied more often to studies with group-comparison design and rarely used with studies with single-case design (Allison & Gorman, 1993; Beeson & Robey, 2006; Miller, 2005; Newell & Burnard, 2006; Van den Noortgate & Onghena, 2003a & 2003b). Furthermore, the results of meta-analysis of single-case studies may not be comparable with the results of meta-analysis of group-comparison studies (Jenson et al., 2007; Parker & Hagan-Burke, 2007). The effect sizes obtained from the meta-analysis of single-case studies are generally larger than the ones from

group-comparison studies (Jenson et al., 2007; Maughan, Christiansen, Jenson, Olympia, & Clark, 2005; Van den Noortgate & Onghena, 2003a).

There are major challenges with the quantification and synthesis of single-case studies. The key challenges result from the specific features of single-case research designs such as different scales being used in different studies and the autocorrelation and trend in the data (Beeson & Robey, 2006; Brossart et al. 2006; Jenson et al., 2007; Miller, 2005). For instance, in order to integrate data across studies, the raw data on the dependent variables (DVs) from different studies, originally measured on different scales, must first be standardized for further comparison (Raudenbush & Bryk, 2002; Van den Noortgate & Onghena, 2003b). In addition to the problem with different scales, the second challenge with quantitative synthesis of single-case research studies is the autocorrelation of data measured at different time points. Autocorrelation is the correlation between the sequential data points that is always assumed to exist in repeated measurements of the same individual over time due to serial dependency of the data. When autocorrelation exists, the assumption of data independence is violated, which may lead to the overestimation or underestimation of Type I errors (Brossart et al., 2006; Jenson et al., 2007; Miller, 2005; Parker & Brossart, 2003; Raudenbush & Bryk, 2002). As well, judgment errors of visual analysis increase when there is autocorrelation in data (Brossart et al., 2006; Manolov & Solanas, 2008). Similarly, a trend in the data (the upward or downward direction of the data point line regardless of intervention) can also result in a bias in the effect size and Type I errors (Miller, 2005). Additionally, Allison and Gorman (1993) indicated that

minor challenges, including disputes on the selection of baseline and intervention phases for comparison (e.g., choosing the first or last intervention phase to compare with the first or last baseline phase) and unavailability of exact values of the data points, could be common problems in the meta-analysis of single-case research.

Several statistical methods have been proposed to provide standardized scales for the synthesis of multiple single-case studies, including mean baseline reduction (MBLR; Campbell, 2004), the percentage of nonoverlapping data points (PND; Scruggs, Matropieri, & Casto, 1987), the percentage of zero data (PZD; Scotti, Evans, Meyer, & Walker, 1991), the improvement rate difference (IRD; Parker, Vannest, & Brown, 2009), standardized mean differences (SMD; Busk & Serlin, 1992), regression methods (e.g., Allison & Gorman, 1993; Faith, Allison, & Gorman, 1996), and Hierarchical Linear Modeling (HLM; Jenson et al., 2007; Raudenbush & Bryk, 2002; Van den Noortgate & Onghena, 2003b & 2007).

The different methods contribute to the synthesis and quantification of the intervention outcomes; however, many of them have been criticized due to their drawbacks (Campbell, 2004; Jenson et al., 2007; Miller, 2005). Campbell (2004) indicated that nonregression methods, such as MBLD, PND, and PZD, are more compatible with visual analysis and can yield quantitative indices for treatment effects through limited or simple calculation procedure when they are compared with regression methods and HLM. However, these methods can lead to confounding results in some circumstances. For instance, PZD cannot be used in calculating the effectiveness of learning a behavior when the behavior is expected

to increase instead of diminishing to zero in the treatment phase. The effectiveness of studies with large variations in targeted behavior such as social skills tends to be underestimated by using PND. Jenson et al. (2007) indicated that PND is limited when analyzing data with outliers, trends, and great variation.

In contrast, regression methods always involve more complicated calculations but are more powerful for reviewing a great number of studies and dealing with the specific issues of data characteristics in single-case research. Campbell (2004) indicated that regression methods were more powerful when dealing with the trend and autocorrelation problems of data in single-case research and less conservative in detecting treatment effects. Nevertheless, the simulation study done by Manolov and Solanas (2008) did not support the idea that regression methods had a better capacity to deal with autocorrelation issues. Their results indicated the analyses using PND and SMD were less affected by the problems of autocorrelation and small number of data points than regression methods. Additional drawbacks of regression methods, similar to the problems in SMD, are the difficulties in the interpretation of their effect sizes (Brossart et al., 2006; Campbell, 2004; Miller, 2005).

In order to solve meta-analysis problems associated with the data from single-case studies, Van den Noortgate and Onghena (2003b, 2007), Raudenbush and Bryk, (2002), and Jenson et al. (2007) have suggested the use of HLM to examine the outcomes across different single-case studies. The scores from HLM are generally structured in units. For instance, the data points from the same dependent variable and the effect sizes from the same participant or the same

study are input into the units using HLM and analyzed as the same category. In the case of meta-analysis of varied studies using different scales, the data from multiple studies are transformed to a standardized metric prior to the HLM analysis, so that the outcomes can be compared on the same scale. A simple HLM equation can be exemplified as follows:

Level-1 Model

$$d_j = \beta_j + e_j$$

Level-2 Model

$$\beta_j = \gamma_0 + \mu_j$$

In the Level-1 model, the observed effect size, d_j, equals the sum of β_j, the true effect size for participant j, and e_j the random error term for the observed effect size from participant j. In the Level-2 model, β_j is the effect size for participant j. γ_0 equals the grand mean effect size across all participants and μ_j is an error term.

By setting hierarchical structure to the data, the researcher can take into consideration the dependence of the scores that is caused by the influence of being in the same unit or category (Miller, 2005; Van den Noortgate & Onghena, 2003b & 2007). With HLM, the researchers can manage the scaling and dependence problems of the scores from single-case studies. Van den Noortgate and Onghena (2007) demonstrated that HLM was superior to regression methods with its strengths in taking autocorrelation into account for studies with a small number of data points. Jenson et al. (2007) supported the application of HLM for the analysis of single-case research due to HLM's powers in management of Type I errors and

detection of treatment effects with average sample sizes. Additionally, researchers can use HLM to describe the variance at the same level of these units by adopting the characteristics of the units as the predictors or moderators. Moderators such as treatment type, sex, and age can be added to the formula in order to examine their effects on intervention outcomes. The application of the HLM to the meta-analysis can help estimate the mean and variation of effect size parameters across multiple studies and test how well different predictors can explain the variation (Miller, 2005; Raudenbush & Bryk, 2002; Van den Noortgate & Onghena, 2007).

Moreover, one great advantage of HLM is its flexibility in modifying the equation according to data characteristics or research interests (Van den Noortgate & Onghena, 2003a). For example, the model can be extended to a three-level one if the meta-analysis includes studies with several dependent variables of varied participants reported in each study. The flexibility of HLM, along with the availability of software programs, further simplifies the calculation procedure for parameter estimations and hypothesis testing.

Because applying HLM to the meta-analysis of single-case research is a new approach, studies that examine the applicability of HLM and compare HLM with other methods are scarce. Proponents (e.g, Miller, 2005; Van den Noortgate & Onghena., 2003b & 2007) support the use of HLM because of its power in dealing with a variety of problems in single-case research data such as different scales across multiple studies, autocorrelation, trend, and small number of data points. However, Van den Noortgate and Onghena (2003a) indicated HLM could be

limited when there is insufficient information for variance analysis, when the data violate assumptions for maximum likelihood estimates, and when there are different conclusions caused by varied models proposed for HLM analysis. More investigation on the strengths and limitations of HLM is required.

In sum, the synthesis of SSIs for individuals with ASD involves the examination of methodological quality and effectiveness of single-case studies. Many meta-analysis methods with different strengths and limitations have been proposed for the synthesis of single-case studies. HLM, a new meta-analysis method, has been proposed as HLM can deal with several specific data issues associated with single-case research, although there are still some limitations with its application. Applying meta-analysis methods to intervention outcomes of single-case research, as a supplement to visual analysis and clinical significance, can provide researchers and clinicians with rich and practical information about treatment effects. Nevertheless, different meta-analysis methods can result in varied outcomes and many factors such as study characteristics, participant characteristics, and data characteristics (i.e., outliers, trends, or number of data points) can impact the results of meta-analysis. Therefore, further studies that investigate the fit between meta-analysis methods and these characteristics are required in order to come up with useful and meaningful information about intervention effectiveness for researchers and practitioners.

Current Dissertation

The present dissertation consists of three papers that examined the methodological quality and treatment effectiveness of single-case SSI studies for

individuals with ASD. The first paper focused on developing a practitioner-friendly quality checklist for single-case research papers. Professionals who are interested in single-case SSI studies for individuals with ASD could use the checklist easily to examine the research quality of the related studies with minimal training in research methodology. A list of quality indicators was recruited from several papers that focus on methodological examination of single-case research or SSIs for individuals with ASD. In this study, an example of how to use the quality indicator checklist was provided by applying the quality checklist to ten single-case studies of SSIs for individuals with ASD. The first study was published in *Developmental Disabilities Bulletin* (Wang & Parrila, 2008).

The second study tested the applicability of hierarchical linear modeling (HLM) for assessing the overall effectiveness of single-case studies using one of two common SSI approaches, peer-mediated and video-modeling methods, with individuals with ASD. Both approaches were found to improve social behavior of individuals with ASD significantly and equally. Furthermore, age was recognized as a significant moderator as younger participants were found to benefit more. The second paper was published in *Research in Autism Spectrum Disorders* (Wang, Cui, & Parrila, 2011).

The third study applied HLM to the examination of the treatment effectiveness of 115 single-case SSI studies for individuals with ASD (Wang, Parrila, & Cui, 2012). The third study also examined what impact different moderators across the three levels of dependent variables, participants, and studies

may have on the treatment effectiveness. The results indicated that the reviewed studies in general yielded positive outcomes. Additionally, one of the moderators examined, the use of the multiple baseline or reversal design, was found to impact treatment outcome positively. All three papers in this dissertation can provide important information for researchers, practitioners, policy-makers, and families in order to choose appropriate SSIs to facilitate the progress of individuals with ASD.

References

Allison, D. B., & Gorman, B. S. (1993). Calculating effect sizes for meta-analyses: The case of the single case. *Behavior Research and Therapy, 31*, 621-631.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington DC: Author.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text revision). Washington DC: Author.

American Psychiatric Association (2012, January 20). DSM-5 Autism Spectrum Disorder News Release 1-20-12. Retrieved from American Psychiatric Association website: http://www.dsm5.org/Documents/12-03%20Autism%20Spectrum%20Disorders%20-%20DSM5.pdf

Apple, A. L., Billingsley, F., & Schwartz, I. S. (2005). Effects of video modeling alone and with self-management on compliment-giving behaviors of children with high-functioning ASD. *Journal of Positive Behavior Interventions, 7,* 33-46.

Baird, G., Simonoff, E., Pickles, A., Chandler, S., Loucas, T., & Meldrum, D. T. (2006). Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames: the Special Needs and Autism Project (SNAP). *Lancet, 368,* 210-215.

Barnhill, G. P., Cook, K. T., Tebbenkamp, K., & Myles, B. S. (2002). The effectiveness of social skills intervention targeting nonverbal communication for adolescents with Asperger syndrome and related

pervasive developmental delays. *Focus on Autism and Other Developmental Disabilities, 17*(2), 112–118.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic-child have a theory of mind. *Cognition, 21,* 37–46.

Barton, M., & Volkmar, F. (1998). How commonly are known medical conditions associated with autism? *Journal of Autism and Developmental Disorders, 28,* 273-278.

Bass, J. D., & Mulick J.A. (2007). Social play skill enhancement of children with autism using peers and siblings as therapists. *Psychology in the School, 44,* 727-734.

Beeson, P. M., & Robey, R. R. (2006). Evaluating single-subject treatment research: Lessons learned from the aphasia literature. *Neuropsychology Review 16,* 161-169.

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,* 264–287.

Brossart, D. F., Parker, R. I., Olson, E. A., & Mahadevan, L. (2006). The relationship between visual analysis and five statistical analyses in a simple AB single-case research design. *Behavior Modification, 30,* 531-563.

Busk, P. L., & Serlin, R C. (1992). Meta-analysis for single-case research. In T. R. Kratochwill & J. R. Levinn (Eds.), *Single-case research design and analysis: New directions for psychology and education* (pp. 187-212). Hillsdale, NJ: Erlbaum.

Campbell, J. M. (2004). Statistical comparison of four effect sizes for single-subject designs. *Behavior Modification, 28,* 234-246.

Carter, A. S., Davis, N. O., Klin, A., & Volkmar, F. R. (2005). Social development in autism. In F. R. Volkmar, R. Paul, A. Klin, & D. A. Cohen (Eds). *Handbook of autism and pervasive developmental disorders* (3rd Edition, pp.312-333). Hoboken, NJ: John Wiley.

Centers for Disease Control (CDC) (2012). *New data on autism spectrum disorders*. Retrieved from http://www.cdc.gov/Features/CountingAutism/

Chakrabarti, S., & Fombonne, E. (2005). Pervasive developmental disorders in preschool children: confirmation of high prevalence. *American Journal of Psychiatry*, *162*, 1133-1141.

Charman, T. (2002). The prevalence of autism spectrum disorders: Recent evidence and future challenges. *European Child & Adolescent Psychiatry, 11*, 249-256.

Cotugno, A. J. (2009). Social competence and social skills training and intervention for children with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *39*, 1268-1277.

Crosbie, J. (1999). Statistical inference in behavior analysis: Useful friend. *Behavior Analyst*, *22*, 105–108.

Curran, S., & Bolton, P. (2009), Genetics of autism. In Kim, Y. (Ed), *Handbook of behavior genetics*, (pp. 397-410). New York, NY, US: Springer Science

Dawson, G. (2008). Early behavioral intervention, brain plasticity, and the prevention of autism spectrum disorder. *Development and Psychopathology,*

20, 775-803.

Faith, M. S., Allison, D. B., & Gorman, B. S. (1996). Meta-analysis of single-case research. In R. D. Franklin, D. B. Allison, & B. S. Gorman (Eds.), *Design and analysis of single-case research* (pp. 245-277). Hillsdale, NJ: Erlbaum.

Fombonne, E. (2005). The changing epidemiology of autism. *Journal of Applied Research in Intellectual Disabilities, 18*, 281-294.

Gillberg, C. (2006). Autism spectrum disorders. In C. Gillberg, R. Harrington, & S. Hans-Charistoph (Eds), *A clinician's handbook of child and adolescent psychiatry.* (pp. 447-488). New York, NY: Cambridge University Press.

Glass, G. (1976).Primary, secondary, and meta-analysis of research. *Educational Researcher, 5*(9), 3-8.

Hall, L. J. (2009). Classification and the physiological approach. In L. J. Hall (Ed), *Autism spectrum disorder from theory to practice* (chap. 1, pp. 1-19). Upper Saddle River, NJ: Person Education, Inc.

Happé, F., & Frith, U. (2006). The weak coherence account: detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders, 36*, 5–25.

Hill, E. L. (2004). Executive dysfunction in autism. *Trends in Cognitive Sciences, 8*, 26–32.

Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children, 71*, 165-179.

Jenson, W. R., Clark, E., Kircher, J. C., & Kristjansson, S. D. (2007) Statistic reform: Evidence-based practice, meta-analyses, and single subject designs. *Psychology in Schools*, *44*, 483-493.

Kasari, C., & Rotheram-Fuller, E. (2007). Peer relationships of children with autism: challenges and interventions. In E. Hollander & E. Anagnostou (Eds.) *Clinical manual for the treatment of autism*. Arlington, VA: American Psychiatric Publishing, Inc.

Koegel, R. L., Werner, G. A., Vismara, L. A., & Koegel, L. K. (2005). The effectiveness of contextually supported play date interactions between children with autism and typically developing peers. *Research and Practice for Persons with Severe Disabilities, 30*, 93-102.

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

Laushey, K. M., & Heflin, L. J. (2000). Enhancing social skills of kindergarten children with autism through the training of multiple peers as tutors. *Journal of Autism and Developmental Disorders, 30*, 183-193.

Liber, D. B., Frea, W. D., & Symon, J. B. G. (2008) Using time-delay to improve social play skills with peers for children with autism. *Journal of Autism and Developmental Disorders, 38*, 312-323.Konstantareas, M. M. (2006). Social skills training in high functioning autism and Asperger's disorder. *Hellenic Journal of Psychology, 3*, 39-56.

Lord, C., Wagner, A., Rogers, S., Szatmari, P., Aman, M., Charman, T., et al. (2005). Challenges in evaluating psychosocial interventions for autistic spectrum disorders. *Journal of Autism and Developmental Disorders, 35,* 695-708.

Manolov, R., & Solanas, A. (2008). Comparing N = 1 effect size indices in presence of autocorrelation. *Behavior Modification, 32,* 860-75.

Marriage, K. J., Gordon, V., & Brand, L. (1995). A social skills group for boys with Asperger's. Australian and New Zealand Journal of Psychiatry, *29,* 58–62.

Matson, J. L., Matson, M. L., & Rivet, T. T. (2007). Social-skills treatments for children with autism spectrum disorders an overview. *Behavior Modification, 31,* 682-707.

Maughan, D. R., Christiansen, E., Jenson, W. R., Olympia, D., & Clark, E. (2005). Behavioral parent training as a treatment for externalizing behaviors and disruptive behavior disorders: A meta-analysis. *School Psychology Review, 34,* 267–286.

McPartland, J., Reichow, B., & Volkmar, F. (2012). Sensitivity and Specificity of Proposed DSM-5 Diagnostic Criteria for Autism Spectrum Disorder *Journal of the American Academy of Child & Adolescent Psychiatry, 51* (4), 368-383.

Miller, A.L. (2005). Interventions targeting reciprocal social interaction in children and young adults with autism spectrum disorders: A meta-analysis. Unpublished doctoral dissertation, University of Utah.

Missouri Autism Guidelines Initiative (2012). *Autism Spectrum Disorders: Guide to Evidence-based Interventions*. Retrieved from *www.autismguidelines.dmh.mo.gov/.../InterventionSummary.pdf*

Morgan, D. L., & Morgan, R. (2009) Data analysis in single-case research. In *Single-case research methods for the behavioral and health sciences* (pp. 205-231). Thousand Oaks, CA: Sage Publications.

National Autism Center (2009). *National standards report: The national standards project—Addressing the need for evidence-based practice guidelines for autism spectrum disorders*. Retrieved from http://www.nationalautismcenter.org/pdf/NAC%20Standards%20Report.pdf

National Institute of Mental Health (2009). *Autistic spectrum disorders pervasive developmental disorders.* Retrieved from National Institute of Mental Health website: http://www.nimh.nih.gov/health/publications/autism/nimhautismspectrum.pdf

National Research Council (NRC). (2001). *Educating children with autism.* Washington, DC: National Academy Press.

Newell, R., & Burnard, P. (2006). *Vital notes for nurses: Research for evidence-based practice*. Oxford, UK: Blackwell.

Parker, R. I., & Brossart, D. F. (2003). Evaluating single-case research data: A comparison of seven statistical methods. *Behavior Therapy, 34*, 189-211.

Parker, R., & Hagan-Burke, S. (2007). Useful effect sizes interpretations for single case research. *Behavior Therapy, 38*, 95-105.

Parker, R. I., Vannest, K. J., & Brown, L. (2009). The improvement rate difference for single-case research. *Exceptional Children, 75*, 135-150.

Rajendran, G., & Mitchell, P. (2007). Cognitive theories of autism. *Developmental Review, 27*, 224-260.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical Linear Models* (2nd ed.). Newbury Park, CA: Sage Publications.

Reichow, B., & Volkmar, F. R. (2010). Social skills interventions for individuals with autism: Evaluation for evidence-based practices within a best evidence synthesis framework. *Journal of Autism and Developmental Disorders*, *40*, 149-166.

Scattone, D. (2007). Social skills intervention for children with autism. *Psychology in the School*, *44*, 717-726.

Scotti, J. R., Evans, L. M., Meyer, L. H., & Walker, P. (1991). A meta-analysis of intervention research with problem behavior: Treatment validity and standards of practice. *American Journal of Mental Retardation, 96,* 233-256.

Scruggs, T. E., Mastropieri, M. A., & Casto, G. (1987). The quantitative synthesis of single-subject research: Methodology and validation. *Remedial and Special Education*, *8*(2), 24-33.

Simpson, R. L. (2005). *Autism spectrum disorders: Intervention and treatments for children and youth* (pp.1-12). Thousand Oaks, CA: Corwin Press.

Smith, T. (2008). Empirically supported and unsupported treatments for autism spectrum disorders. *The Scientific Review of Mental Health Practice, .6,*

3-20.

Smith, T., Scahill, L., Dawson, G., Guthrie, D., Lord, C., Odom, S., Rogers, S., & Wagner, A. (2007). Designing research studies on psychosocial interventions in autism. *Journal of Autism and Developmental Disorders, 37*, 354-366.

Spector, S. G., & Volkmar, F. R. (2006). Autism Spectrum Disorders. In D. A. Wolfe, & E. J. Mash (Eds), *Behavioral and emotional disorders in adolescents: Nature, assessment, and treatment* (pp. 444-460). New York, NY: Guilford Publications.

Strain, P. S., Schwartz, I. S., & Bovey, E. H. (2008). Social competence interventions for young children with autism. In W.H. Brown, S.L. Odom, S.R. McConnell (Eds.), *Social competence of young children: Risk, Disability, & Intervention* (pp. 253-272). Baltimore: Paul H. Brookes.

Van den Noortgate, W., & Onghena, P. (2003a). Combining single-case experimental data using hierarchical linear models. *School Psychology Quarterly, 18*, 325-346.

Van den Noortgate, W., & Onghena, P. (2003b). Hierarchical linear models for the quantitative integration of effect sizes in single-case research. *Behavior Research Methods, Instruments, & Computers, 35*, 1-10.

Van den Noortgate, W., & Onghena, P. (2007). The aggregation of single-case results using hierarchical linear models. *The Behavior Analyst Today, 8*, 196-208.

Volkmar, F. R., & Klin, A. (2005). Issue in the classification of autism and related

conditions. In F. R. Volkmar, A. Klin, R. Paul, & J. D. Cohen (Eds), *Handbook of autism and pervasive developmental disorders* (3rd Edition, pp. 5-41). New York: Wiley.

Volkmar F. R., Lord C., Bailey A, Schultz R. T., & Klin, A. (2004). Autism and pervasive developmental disorders. *Journal of Child Psychology and Psychiatry, 45,* 135–170.

Wang, S.Y., Cui, Y., & Parrila, R. (2011). Examining the effectiveness of peer-mediated and video-modeling social skill interventions for children with autism spectrum disorders: A meta-analysis of single-case research Using HLM. *Research in Autism Spectrum Disorders, 5,* 562-569.

Wang, S.Y., & Parrila, R. (2008). Quality Indicators for Single–Case Research on Social Skill Interventions for Children with Autistic Spectrum Disorder. *Developmental Disabilities Bulletin, 36,* 81-105.

Wang, S. Y., Parrila, R. & Cui, Y. (2012). Meta-Analysis of Social Skills Interventions of Single-Case Research for Individuals with Autism Spectrum Disorders: Results from Three-Level HLM. *Journal of Autism and Developmental Disorders.* Advance online publication. doi: 10.1007/s10803-012-1726-2

White, W. S., Keonig, K., & Scahill, L. (2007). Social skills development in children with autism spectrum disorders: A review of the intervention research. *Journal of Autism and Developmental Disorders, 37,* 1858-1868.

Wing, L. (1991). The relationship between Asperger's syndrome and Kanner's autism. In U. Frith (Ed.), *Autism and Asperger syndrome* (pp. 93-121).

Cambridge, UK: Cambridge University Press.

Wolff, S. (2004). The history of autism. *European Child and Adolescent Psychiatry, 13*, 201-208.

World Health Organization. (1992). *International classification of diseases: Diagnostic criteria for research* (10th ed.). Geneva, Switzerland: Author.

CHAPTER II

QUALITY INDICATORS FOR SINGLE-CASE RESEARCH ON SOCIAL SKILL INTERVENTIONS FOR CHILDREN WITH AUTISTIC SPECTRUM DISORDER

Introduction

Social interaction problems are recognized as one of the core deficits for children with autistic spectrum disorder (ASD; White, Keonig, & Scahill, 2007). Recently, the diagnostic criteria and conceptualization of autistic disorder has been broadened from autism to ASD (Fombonne, 2005), resulting in an increase of individuals with the central deficit in social interaction, such as many with pervasive developmental disorders not otherwise specified (PDDNOS), or Asperger's disorder. Children with high function autism (HFA), PDDNOS, or Asperger's disorder show fewer cognitive and language deficits, but social interaction issues can be a major barrier for them that impacts negatively on their adjustment at school and community. When children with ASD are placed in inclusive settings, they tend to be isolated or experience difficulties in establishing friendships with peers. Even for the children with PDDNOS, Asperger's Syndrome, or HFA, the ones with more preserved cognitive skills, interacting appropriately with others can be a difficult task (Rao, Beidel, & Murray, 2008). Traditional intervention models that stress academic or basic living skills fail to meet their needs and, as a result, interventions targeting social skills have started to gain in popularity.

Several models have been developed for the social skill training of children with ASD, including behavior modification, peer-mediated training,

social story, video modeling, self-management, pivotal response training, joint attention training, and buddy system (e.g., Bass & Mulick, 2007; Matson, Matson, & Rivet, 2007; Scattone, 2007). Although the strategy of modeling target behaviours and providing reinforcement tends to be the most commonly used, there are increasing varieties of the intervention approaches in the newer studies. While some approaches, such as the peer-mediated approach (Matson et al., 2007), have generally proved to be effective for children with ASD, others have produced less consistent findings across different studies.

As more studies on social skill interventions are completed, there is a growing need to assess and integrate the evidence of the efficacy or effectiveness of the interventions they provide. Parents want to know how to choose an effective model for their children with ASD, clinicians would like to adopt the most effective model for their evidence-based practice, and policy-makers are interested in funding programs with proven effectiveness. Therefore, how to examine the quality of the intervention research systematically has become a critical issue for those who are interested in social skill interventions for children with ASD.

This paper describes development of a quality checklist that parents, teachers, clinicians, and policy-makers with basic research skills can use to systematically evaluate the methodological quality of single-case studies on social skill training of children with ASD. We focus on single-case studies because a recent review by Matson et al. (2007) indicated that more than 90% of the intervention studies employ this design. Below, we will describe in more detail the

development of the quality checklist and provide an explanation of the items included. We will also provide two examples of how the checklist can be used, first to examine the overall quality of individual studies, and then to examine the quality of a small body of research. The complete checklist is provided in Appendix A.

Quality Indicators

The first step involved identifying quality indicators that could be used to assess methodological quality of single-case studies. The initial list of the quality indicators was adapted from Horner, Carr, Halle, McGee, Odom, and Wolery (2005). Horner et al. list multiple criteria that can be used to examine different dimensions of single-case research, including information given on participants, settings, dependent and independent variables, baseline data collection, experimental control/internal validity, external validity, and social validity. As Horner et al.'s indicators were not specific to social skills interventions for children with ASD, their criteria were compared to those used in three recent papers that focused more closely on this specific topic (Lord et al., 2005; Smith et al., 2007; Reichow, Volkmar, & Cicchetti, 2008). Several quality indicators were added to the checklist, such as use of standardized instruments for diagnosis, information of the peers and interventionists, the criteria for the percentage of the sessions used to examine the inter-rater agreement, utilization of multiple baseline or reversal design, the amount of data points in the baseline and intervention phases, and the use of blind agents for establishing social validity. Table 2.1 lists the initial quality indicators.

The quality indicators are divided into two parts: primary and secondary quality indicators. The primary quality indicators focus on the internal validity of the research. The more the studies meet the criteria for the primary quality indicators, the better the studies manage the confounding factors and can demonstrate the causal relationship between the intervention and the observed outcomes. We designate internal validity indicators as primary as a study without internal validity cannot produce useful information. The secondary indicators are concerned with the external and social validity of the studies. A study earns credit in external validity when it presents evidence for generalizability of the results across various time frames, settings, or participants. A study demonstrates better social validity if more people recognize the importance of the intervention or give credit to the outcome of the intervention. A more detailed description of the initial quality indicators together with a rationale for their inclusion follows.

Primary quality indicators.

The primary quality indicators are used mainly to examine whether (a) the study includes sufficient information about the participants, settings/material, and the independent and dependent variables, (b) the researchers manipulated and measured variables faithfully, (c) there were sufficient data points and stable pattern in the data, and (d) the study adopted one of the designs that can demonstrate functional relationship between target behavior and intervention.

Participants.

The researcher should provide detailed and precise information regarding the participants. This information is necessary for others to be able to replicate the

research, or to apply the results to different groups of children. Detailed information should be provided about the children with ASD, interventionists, and the peers and parents, if applicable.

First, the information about the children with ASD should include gender, age, ethnicity, recruiting procedure, selection criteria, information on relevant ability such as IQ, academic ability, or adaptive skills, and confirmative information of ASD diagnosis. Age and gender is the basic information in order to facilitate the selection of participants for replication (Horner et al., 2005). Furthermore, age has been found to be related to the differential gains from the intervention, with younger children tending to gain more from the intervention than older children (Baker-Ericzen, Stahmer, & Burns, 2007; Corsello, 2005).

Ethnicity information can demonstrate the demographic characteristic of the population and can be related particularly to the effectiveness of parent education interventions (Baker-Ericzen et al., 2007). Selection criteria and recruiting procedure provide explicit standards regarding what kinds of characteristics the participants exhibited and how they were selected. In addition, the information on relevant abilities such as IQ, language abilities, or the index of social interaction for ASD participants should be provided in details. Because ASD represents a heterogeneous group and the abilities of children within subgroups of ASD can be diverse (Fombonne, 2005; NRC, 2001; White et al., 2007), children with different levels of abilities can respond to the same intervention differently (Shea, 2004; Sherer & Schreibman, 2005). The levels of language ability for different subgroups of ASD can range from no speech to

fluent but idiosyncratic communication (NRC, 2001). The social interaction of children with ASD can be categorized as aloof, passive, or active but odd (Wing & Gould, 1979). The intelligent levels of the children with ASD can range from severe mental retardation to superior levels. NRC (2001) indicates that there likely is no single intervention approach that benefits all different types of children with ASD equally. Thus, detailed information about the abilities of participants with ASD is necessary for both replication and assessment of generalizability.

Similar to the ability levels, the accuracy of ASD diagnosis can interfere greatly with the efficacy of the intervention. Smith et al. (2007) suggested that researchers should ensure faithfulness of ASD diagnosis by using standardized diagnostic tools. Hence, use of diagnostic tools with standardized procedure or description such as CARS (Childhood Autism Rating Scale), ADOS (Autism Diagnostic Observation Schedule), ADI-R (Autism Diagnostic Interview-Revised), DSM -IV, or ICD-10 (International Statistical Classification of Diseases and Related Health Problems-10th version) is included as one of the quality indicators. In addition, diagnosis from psychologists, psychiatrists, or pediatricians is included as a separate quality indicator as it can support the accuracy of the diagnosis in addition to the use of standardized diagnostic tools.

The research on social skill training usually involves interventionists to implement the training and the information regarding the background and training experience of the interventionists should be provided. Replicating research with insufficiently trained interventionists can result in insignificant outcomes and

undermine the efficacy of the intervention model. Furthermore, if the peers or parents participated as mediators in the research (i.e., peer-mediated or parent-mediated model), the researchers should present information such as their recruiting procedure and selection criteria to facilitate future replication and meta-analyses.

Settings and materials used for social skill training.

The information on settings and materials is important as different settings and materials may motivate children differently and interfere with their social interaction dramatically even without intervention. Therefore, the researcher should provide sufficient information regarding how the setting was arranged or what type of materials – such as games or tools – were available and used. Structuring the setting and materials in a consistent way can demonstrate the functional relationship between the outcomes and intervention more clearly.

Independent variables (IV).

In social skill training, independent variables (IVs) are the specific procedures or strategies used for intervention, and the implementation of IVs should lead to the change of the social behavior. Clear and detailed descriptions of independent variables (IV) are necessary for replication and generalization studies. Using standardized manuals for implementing IVs generally ensures there is detailed information to repeat the procedure, and the manual also can be used for creating a checklist to examine if the intervention is being implemented faithfully.

Furthermore, a study can earn credit if it tries to control possible confounding factors (i.e., contamination effects between children), manipulate IVs

at least three different times (Reichow et al., 2008), and assess the fidelity of IV implementation. These quality indicators examine whether the researchers have provided sufficient evidence to support the linkage between the observed behaviors and the IV.

Dependent variables (DV).

Dependent variables (DV) are the measurements of the target behaviors that the researchers aim to change (either increase or decrease) with the implementation of IVs. All possible target behaviors need to be defined operationally so that they can be measured with minimal error, and the measured behaviors have to be clearly connected to socially desired outcomes that they are chosen to represent. Finally, the measurement procedure has to be clearly described to allow replication.

In order to demonstrate the effect of intervention, data on DVs should be collected a minimum of three times during each baseline and intervention phase (Horner et al., 2005; Reichow et al., 2008). Further, data should display a stable pattern or trend at each phase. Without stable pattern or trend, the study cannot provide sufficient evidence for the differences between the phases. Lack of stable pattern or trend may also indicate presence of confounding factors. As a result, verifying a fundamental link between IV and DV by contrasting the patterns at different phases becomes difficult.

In addition, because most measurements of DVs in social skill intervention studies involve raters, the researchers should test the reliability of the measurement by comparing the rating outcomes across different raters for a

minimum of 20% of the sessions (Reichow et al., 2008), and the inter-rater agreement should be at least 0.6 or Kappa coefficient over .60 (Horner et al., 2005; Reichow et al., 2008). Moreover, if the study includes raters that are different from interventionists and raters are blind to the research, validity of the ratings is further increased.

Research designs.

An additional quality indicator was added to indicate whether the study used a design that clearly can support a functional relationship between targeted social behaviors and the intervention. Multiple-baseline and reversal designs were chosen as preferred designs because both decrease threats to internal validity and provide a more powerful statement for the efficacy of the intervention (NRC, 2001; Richards, Taylor, Ramasamy, & Richards, 1999; Smith et al., 2007). In the multiple-baseline design, the researchers implement the intervention to different participants at different settings, or to different behaviors at different time frames. If the change in dependent variables corresponds to the implementation of the intervention at different time points across different participants, settings, or behaviors, more convincing evidence to support the effect of the intervention is generated. With the use of reversal design, the study can rule out the effect of maturation and history that generally confound the interpretation of the intervention effect in a simple A-B design. The reversal design also provides opportunities to examine the generalization effect of the intervention. The alternating design is not appropriate for examining intervention outcomes of social skill training because the effect of one intervention can interfere with the

possible effect of the other intervention. In addition, the changing criterion design is aimed to increase or decrease developed skills, and changing criterion design may not be appropriate in social skill training because social skill training usually involves developing new skills.

Secondary quality indicators.

The secondary quality indicators are used to examine external and social validity of the research. External validity is mainly concerned with the generalizability of the results to different settings and participants, whereas social validity is mainly concerned with recognized social importance of the intervention outcomes.

External validity.

External validity is regarded as high when the target behavior is maintained over longer periods of time and we have a reason a believe that the positive effect of intervention can be generalized to different individuals in different contexts; after all, the ultimate goal of the intervention research is to find interventions that benefit more participants with similar difficulties and maintain the gains across different settings and time. The researchers can assess maintenance effects of the intervention by measuring the DVs again some time after the intervention has been discontinued. In order to distinguish maintenance effect from generalization effect, the data on maintenance effect should be obtained with the presence of the same experimental setting, participants, and materials as was used during the intervention. The generalized effects of the intervention can be assessed through measuring DVs while having ASD

participants interact with different persons, or with the same persons in different settings, or with the same person but with different activities or toys. Hence, the external validity of the study is increased if the researchers included maintenance or follow-up sessions over an extended period, and if they verify the effect of the intervention in different contexts.

Social validity.

Social validity is increased in a single-case study if the study examines social importance of the intervention outcomes to the children with ASD and to other people around the children, such as school staff, teachers, friends, siblings or parents, or society. Social validity can be established directly or indirectly. For instance, children, parents, or teachers may report how well the intervention had improved children's social interactions other than the DVs that the researchers measured. In other cases, the intervention may indirectly benefit children's self-esteem or child-parent relationship that were not the primary focus of the intervention. Therefore, social validity indicators include measurements of the direct and secondary gains of the intervention, consumer satisfaction reports, and qualitative reports of the progress. The direct gains are related to the improvement of social behavior other than the DVs, whereas the secondary gains are defined as the improvement of non-social behavior or psychological status such as child-parent relationship, self-value, self-confidence, happiness, disruptive behaviors or social alliance. Moreover, if the implementation of IV is conducted in a context close to natural settings, there will be better chances for adaptation to real world settings. As a result, the social validity of the study is strengthened if

minimal adjustment of IV implementation is required for real world settings.

In addition, using raters that are blind to the research to evaluate social validity is counted as one of the quality indicators because blindness of raters will decrease the possible confirmation bias. Additional credit will be given if the researchers examine social validity three months or longer after the intervention has stopped. Furthermore, another quality indicator, based on a suggestion by Horner et al. (2005), is placed in the checklist to inspect whether the IV implementation is cost-effective and time-effective.

Examples of How the Quality Indicator Checklist Can Be Used

After the initial checklist of quality indicators was developed, we used the checklist to examine research papers that reported single-case research with focus on the social skill training of children with ASD. The quality indicators in the initial list used to examine the cost-effectiveness and time-effectiveness of the implementation were excluded after probing the first two papers. The reason for exclusion was the lack of agreement as to the operational definitions for cost-effectiveness or time-effectiveness. The revised quality checklist had 39 quality indicators remaining (see Appendix A).

Target papers.

Thirty recent (published between 2000 and 2007) papers were located either through Academic Search Premier, Web of Science, and TOC Premier databases using keywords “autism,” “social skill,” “intervention” and “training,” or from the reference lists of recent review articles on social skills interventions for children with ASD (Bass et al., 2007; Matson et al., 2007; Rao et al., 2008;

Scattone, 2007). Five papers were excluded because they did not meet most criteria of single-case studies. Ten of the remaining 25 papers that adopted one or more models of behavior modification, peer-mediated training, social story, pivotal response training, joint attention training, or buddy system were randomly selected for this review. The 10 papers are summarized in Table 2.2 and numbered by superscript in the reference list. The total number of children with ASD was 28 including two females and 26 males. The ages of children with ASD ranged from three to nine years-old. The number of participants with ASD within these studies ranged from one to five. The intervention models or strategies used in these studies included behavior analysis, pivotal response training, peer-mediated approach, social story, role play/modeling/prompt/prime/reinforcement, and social script; 70 % of them used more than one intervention model or strategy. Two studies did not report the duration or frequency of the intervention because they used varied cut-off criteria for different phrases of intervention and the duration or frequency of the intervention for each participant with ASD varied. Two studies did not do so because the duration or intensity of the intervention could not be accumulated due to the adoption of all classmates or the context as independent variables. Four of the ten studies indicated significant improvement across all target behaviors, while the remaining studies indicated that some target behaviors improved or that some of the ASD participants showed improvement in social behaviors. In terms of settings, four of the studies were conducted in a public school, two were conducted in laboratory settings, two in a private school or private education center, one in a community clinic, and one in children's

community or their homes.

Scoring of individual papers.

Thirty-nine remaining quality indicators were used to examine the ten papers. If the paper met the criterion for a specific quality indicator, it was given one point for that item. If partial criterion was met, 0.5 point was given. If the item was not applicable (for example, the quality indicator for detailed information about peers is not applicable to the studies that don't involve peers for intervention), it was counted as "not applicable." Thus, the maximum score varied across studies and was less than 39 for the studies for whom not all quality indicators could be applied. To provide a common metric across studies, we calculated the proportion of applicable quality indicators met. For example, if the study received 25 points across 38 applicable items, its total score was 25/38 = .66. The total score results are presented on the last line of Table 2.2 and can be construed as representing an assessment of the overall quality of the paper. Note, however, that the total scores are somewhat simplistic estimates of the total quality of the studies as all quality indicators were given equal weighting. Similar ratio scores were also calculated separately for the primary quality indicators and the secondary quality indicators.

The total quality scores of the reviewed papers ranged from .33 to .65 with the mean of .52 (*SD* = 0.10) indicating that, on average, these studies met about half of the applicable quality indicators. The scores of the 10 papers ranged from .28 to .74 over the twenty-nine primary indicators with the mean of .59 (*SD* = 0.14). The scores of the 10 papers ranged from 0 to .5 over the ten secondary

indicators with the mean of .33 (*SD* = 0.14).

Compared with other papers, the paper with the highest total score met most primary quality indicators, with the exception of the indicators of using standardized procedure for implementation, providing detailed ethnicity information of ASD participants, providing detailed information regarding the training or qualification of interventionist, demonstrating stable patterns in baseline and intervention phases, and having the interventionist different from experimenters or blind to research. However, the study demonstrated the functional relationship by adopting multiple-baseline design.

Examination of the results by quality indicators.

Table 2.3 presents the results across different quality indicators and can be used to examine the overall quality of this small body of research and to identify specific problems that may be replicated across multiple studies. On the positive side, Table 2.3 shows that all of the ten papers provided information of their participants' gender and age, manipulated IV at least three different times, provided an operational definition of DV, linked DV as measured clearly to the target behaviors, generated quantifiable index for DV, and repeated measurement at least 3 times at each intervention phase. Seven to nine papers also provided detailed information of recruiting procedures for peers, described the IV in detail, measured DV at least 3 times at baseline phase, reached 80 % interrater agreement or 0.6 kappa index, adopted either multiple baseline or reversal design, and used interventions that require minimal adjustments for implementation in natural settings. Six papers provided detailed information about the selection criteria for

the ASD children, used standardized instruments for diagnosis, and collected data on maintenance effects, and three out of five papers that used peers included information on their selection criteria.

The criteria that half or more than half of the papers did not meet included providing information on ethnicity and relevant abilities of the ASD participants, as well as whether they were diagnosed with ASD by professionals specialized in autism. Half of the papers reported data on generalization of the effects to different contexts. Only three to four papers provided detailed information on the training and qualifications of the interventionists, or the settings and materials that were used. Four papers controlled for materials and settings, contamination between subjects, fidelity of implementation, and rater-bias. While most papers collected sufficient amounts of data, only three showed a stable pattern/trend in baseline phase, and only one paper showed a stable pattern/trend on each intervention phase.

Finally, most papers fared poorly in terms of the social validity quality indicators, indicating that this is an area where there is ample room for improvement.

Discussion

Interacting appropriately with others is a significant challenge to many children with ASD and intervention studies targeting social skills have increased both in popularity and in variety. Several models have been developed and tested for the social skill training of children with ASD, and both the outcomes and the quality of the studies evaluating the models vary widely. With more studies

published weekly, there is a growing need for tools that help not only researchers but also parents, teachers, clinicians, and policy-makers to assess the accumulating evidence for different models. One important part of this assessment is the examination of the quality of research used to support different intervention programs; only high-quality studies can provide a basis for evidence-based practice, and choosing an intervention program or programs to implement and fund requires examination of both the effectiveness of those programs as well as the quality of the studies establishing the effectiveness. How to examine the quality of the intervention research systematically has become a critical issue for those who are interested in social skill interventions for children with ASD. Hence, this paper aimed to develop a checklist of quality indicators that can be used by a variety of people with basic research skills to systematically review the quality single-case studies of social skill intervention for children with ASD.

The developed checklist includes several quality indicators for examining internal, external and social validities of the single-case research papers. Parents, teachers, clinicians, and policy-makers with basic research skills can go through and check the criteria of the checklist one by one while reading each research paper. They can give credits to the study for its internal validity by examining whether there is detailed information of participants, interventionist, IVs, and DVs, sufficient and reliable data-points across phases, control over confounding factors, and a research design that can demonstrate a functional relationship between the intervention and the outcome. In particular, providing sufficient information on different aspects of the study is important because it allows replications of the

studies that are necessary for establishing the efficacy of any intervention. External validity is established if the study applies the intervention to different interactive people, settings, or materials. Furthermore, the study can earn credits on social validity when it reports on how the participants and other people recognized the contribution of the intervention. However, the indicators of internal validity are more important when judging the overall quality of the study than the indicators of external validity or social validity. Simply put, if the study lacks internal validity, there are no valid results that can be generalized or proven socially important. To acknowledge this, we clustered the indicators related to internal validity under the heading of primary quality indicators.

In the second half of the paper, we provided examples of how the quality indicator checklist can be used to assess both the quality of individual studies as well as the quality of a body of research on a specific topic. Taken together, Tables 2.2 and 2.3 indicate significant flaws in both internal and external validity indicators, and that no single study is clearly above the criticism. For example, none of the papers in this review provided all of the information needed for a replication study, and only one included most required information (except the information regarding the training background or qualifications of the interventionist). In addition, many of those studies could improve their quality if they have had examined the fidelity of the implementation or provided operational definitions and measurable indexes for both the IV and the DV. Although the ten papers selected for the review may not fully represent the field, the results of this review highlight the need to examine the quality of studies carefully before

accepting their results. Furthermore, researchers interested in studying social skill intervention programs for children with ASD would benefit from using this quality checklist to examine how well they have designed and reported their studies.

Some limitations with the quality indicators should be noted. First, the quality indicator checklist may require additional modification when it is used to examine a large body of papers. For example, the now excluded criteria of cost- and time-effectiveness could be added back if proper operational definitions become available. Those criteria can be important when we try to examine if the intervention model can be implemented in natural settings. In addition, using the total scores to rank the studies should be done with care because each indicator is now given equal weight. The primary indicators should be weighted more in the final decision since those items are central to the quality of the research.

Footnote. A version of this chapter has been published. Wang & Parrila 2008. Developmental Disabilities Bulletin. 36: 81-105.

Table 2.1.

The Initial Checklist of Quality Indicators for Single-Case Studies of Social Skills Training for Children with ASD

Primary Quality indicators

Participants:

Gender and age of ASD participant is provided

Ethnicity information of ASD participant is provided

Recruiting procedure of ASD participant is explained

Selection criteria of ASD participant are explained

IQ, academic performance, or adaptive skills data provided

ASD diagnosis made by professionals specialized in autism

The study used a standardized instrument for diagnosis

Detailed information on training & qualifications of interventionists provided

Detailed information on the recruiting procedure of peers provided (if applicable)

Detailed information of selection criteria of peers (if applicable) provided

Settings/materials used for social skill training:

Information of the settings and materials sufficient for replication

Potential confounding factors caused by the settings/materials controlled

Independent Variables:

IVs were described in sufficient detail for replication

Standardized procedure used for implementation (i.e., manual)

IV implemented at least three times at three different time points

Researchers controlled the contamination between subjects

The researchers assessed the fidelity of implementation

Dependent Variables:

DVs were operationally defined

DV is clearly linked to target behaviors

DVs measured at least 3 times on each baseline phase

Measuring procedure generated a quantifiable index

The data on each baseline phase presents a stable pattern/trend

DVs measured at least 3 times on each intervention phase

The data on each intervention phase present a stable pattern/trend

The inter-rater agreement is over 80% or Kappa over .60 between raters

The inter-rater agreement was collected on at least 20% of sessions

The raters were blind to research

The raters were different from the interventionist

Research Design: The study used multiple baseline or reversal design

Secondary Quality indicators

External Validity:

The researcher reported data on maintenance effect

The data on generalization of effects are collected across different contexts

Social Validity:

Data on direct gains (other than DVs) caused by intervention reported

Data on secondary gains caused by intervention reported

Data on consumer satisfaction reported

Qualitative data reported for social importance of change in DVs

The implementation of IV cost- and time-effective

IV implementation needs minimal adjustment to natural settings

The research examined SV over extended (3 month later) period

The agents used to establish SV blind to research

The agents used to establish SV adopted from typical contexts

Table 2.2.

Summary of the Reviewed Papers

Study ID	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Publication year	2005	2007	2003	2006	2000	2003	2008	2007	2008	2005
Settings	Priv	Lab	Clin	P.S.	P.S.	Lab	P.S.	P.S.	Priv	home/ community
Number of children with ASD	1	4	4	3	2	5	2	2	3	2
Number of Male	1	4	3	3	2	5	2	2	3	1
Number of Female	0	0	1	0	0	0	0	0	0	1
Design	M-B	AB	AB	M-B	reversal	M-B	M-B	M-B	M-B	M-B
Models/ strategies										
Behavior Analysis	☆					☆			☆	
pivotal response training						☆	☆			☆
peer mediated		☆		☆			☆	☆		
social story				☆						
social script			☆							
role play/modeling/prompt/ prime/ reinforcement	☆	☆	☆	☆	☆	☆	☆		☆	
outcome of intervention	★	★	★	★	★★	★	★	★★	★★	★★
generalization of intervention	★	NA	★	★	★★	★	★	★	★★	★★
Score for Primary QI	.28	.48	.62	.53	.67	.74	.64	.72	.53	.66
Score for Secondary QI	.50	.00	.40	.38	.25	.40	.30	.40	.30	.40
TOTAL SCORE	.33	.38	.56	.50	.57	.65	.55	.64	.47	.59

NOTE: Study ID identifies the study in question in the reference list; Priv = private school or educational center; Lab = laboratory; Clin = clinic; P.S. = public school; M-B: multiple-baseline design; AB= one baseline session + one intervention session; Reversal = design includes withdrawal phase; ★ = indicate partial improvement in the outcome of target behaviors; ★★ = indicate improvement in all target behaviors; NA = not applicable.

Table 2.3.
The percentage of papers meeting the criteria of each primary and secondary quality indicators

Quality indicators	**%**			
Primary Quality indicators	Yes	No	Part	NA
Participants:				
Gender and age of ASD participant is provided	100	0	0	0
Ethnicity information of ASD participant is provided	40	60	0	0
Recruiting procedure of ASD participant is explained	30	70	0	0
Selection criteria of ASD participant are explained	60	40	0	0
Information of relevant abilities (IQ, academic performance, or adaptive skills) provided	40	60	0	0
ASD diagnosis made by professionals specialized in autism	20	80	0	0
The study used a standardized instrument for diagnosis	60	40	0	0
Detailed information on training & qualifications of interventionists provided	30	70	0	0
Detailed information on the recruiting procedure of peers provided	70	10	0	20
Detailed information of selection criteria of peers provided	30	50	0	20
Settings/materials used for social skill training:				
Information of the settings and materials sufficient for replication	40	60	0	0
Potential confounding factors caused by the settings/materials controlled	40	60	0	0
Independent Variables:				
IVs described in sufficient detail for replication	70	30	0	0
Standardized procedure used for implementation (i.e., manual)	30	70	0	0
Researchers controlled the contamination between subjects	40	60	0	0
IV implemented at least three times at three different time points	100	0	0	0
The researchers assessed the fidelity of implementation	40	60	0	0
Dependent Variables:				
DVs were operationally defined	100	0	0	0
DV is clearly linked to target behaviors	100	0	0	0
Measuring procedure generated a quantifiable index	100	0	0	0
DVs measured at least 3 times on each baseline phase	80	20	0	0
The data on each baseline phase presents a stable pattern/trend	30	40	30	0
DVs measured at least 3 times on each intervention phase	100	0	0	0
The data on each intervention phase present a stable pattern/trend	10	80	10	0

(Table 2.3 continues)

(Table 2.3 continued)

The inter-rater agreement is over 80% or Kappa over .60 between raters	90	0	10	0
The inter-rater agreement was collected on at least 20% of session	90	10	0	0
The raters were blind to research	0	100	0	0
The raters were different from the interventionist	40	60	0	0
Research Designs: using multiple baseline or reversal design	70	30	0	0
Secondary Quality indicators				
External validity:				
The researcher reported data on maintenance effect	60	40	0	0
The data on generalization of effects are collected across different contexts	50	50	0	0
Social validity:				
Data on direct gains (other than DVs) caused by intervention reported	40	60	0	0
Data on secondary gains caused by intervention reported	20	80	0	0
Data on consumer satisfaction reported	10	90	0	0
Qualitative data reported for social importance of change in DVs	20	80	0	0
IV implementation needs minimal adjustment to natural settings	70	30	0	0
The research examined SV over extended (3 month later) period	0	100	0	0
The agents used to establish SV blind to research	10	60	0	30
The agents used to establish SV adopted from typical contexts	40	30	0	30

Note: Total numbers of reviewed papers=10; Yes = Meet the criterion; No = Do not meet the criterion; Part = Meet the criterion partially; NA = The criterion not applied to the paper

References

Baker-Ericzen, M.J., Stahmer, A.C., & Burns, A. (2007). Child demographics associated with outcomes in a community-based pivotal response training program. *Journal of Positive Behavior Interventions, 9*, 52–60.

(3) Barry, T. D., Klinger, L. G., Lee, J. M., Palardy, N., Gilmore, T., & Bodin, S.D. (2003). Examining the effectiveness of an outpatient clinic-based social skills group for high-functioning children with autism. *Journal of Autism and Developmental Disorders, 33*, 685-701.

Bass, J. D., & Mulick J.A. (2007). Social play skill enhancement of children with autism using peers and siblings as therapists. *Psychology in the School, 44*, 727-734.

(2) Chung, K.M., Reavis, S. Mosconi, M., Drewry, J., Matthews, J., & Tasse, M. J. (2007). Peer-mediated social skills training program for young children with high-functioning autism. *Research in Developmental Disabilities, 28,* 423-436.

Corsello, C.M. (2005). Early intervention in autism. *Infant and Young Children, 18*, 74-88.

(4) Delano, M., & Snell, M. E. (2006). The effects of social stories on the social engagement of children with autism. *Journal of Positive Behavior Interventions, 8,* 29-42.

Fombonne, E. (2005). The changing epidemiology of autism. *Journal of Applied Research in Intellectual Disabilities, 18,* 281-294.

(7) Harper, C. B., Symon J. B. G., & Frea, W. D. (2008). Recess is time-in: using

peers to improve social skills of children with autism. *Journal of Autism and Developmental Disorders, 38,* 815-826.

Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children, 71,* 165-179.

(10) Koegel, R. L., Werner, G. A., Vismara, L. A., & Koegel, L. K. (2005). The effectiveness of contextually supported play date interactions between children with autism and typically developing peers. *Research and Practice for Persons with Severe Disabilities, 30,* 93-102.

(5) Laushey, K. M., & Heflin, L. J. (2000). Enhancing social skills of kindergarten children with autism through the training of multiple peers as tutors. *Journal of Autism and Developmental Disorders, 30,* 183-193.

(8) Lee, S., Odom, S. L., & Loftin, R. (2007). Social engagement with peers and stereotypic behavior of children with autism. *Journal of Positive Behavior Interventions, 9,* 67-79.

(9) Liber, D. B., Frea, W. D., & Symon, J. B. G. (2008) Using time-delay to improve social play skills with peers for children with autism. *Journal of Autism and Developmental Disorders, 38,* 312-323.

Lord, C., Wagner, A., Rogers, S., Szatmari, P., Aman, M., Charman, T., et al. (2005). Challenges in evaluating psychosocial interventions for autistic spectrum disorders. *Journal of Autism and Developmental Disorders, 35,* 695-708.

Matson, J. L., Matson, M. L., & Rivet, T. T. (2007). Social-skills treatments for

children with autism spectrum disorders an overview. *Behavior Modification, 31*, 682-707.

National Research Council (NRC). (2001). *Educating children with autism,* National Academy Press, Washington, DC.

Rao, P. A., Beidel, D. C., & Murray, M. J. (2008). Social skills interventions for children with Asperger's syndrome or high-functioning autism: a review and recommendations. *Journal of Autism and Developmental Disorders, 38*, 353-361.

Reichow, B., Volkmar, F.R., & Cicchetti, D.V. (2008). Development of the evaluative method for evaluating and determining evidence-based practices in autism. *Journal of Autism and Developmental Disorders, 38*, 1311-1319.

Richards, S. B., Taylor, R. L., Ramasamy, R., & Richards, R. Y. (1999). Single subject research application in educational and clinical settings. (pp.134-147 & 145-164). San Diego: Singular Publishing Groups, Inc..

(1) Sawyer, L. M., Luiselli, J. M., Riciardi, J. N., & Gower, J. L. (2005). Teaching a child with autism to share among peers in an integrated preschool classroom: Acquisition, maintenance, and social validity. *Education and Treatment of Children, 28*, 1-10.

Scattone, D. (2007). Social skills intervention for children with autism. *Psychology in the School, 44*, 717-726.

Shea, V. (2004). A perspective on the research literature related to early intensive behavioral intervention (Lovaas) for young children with autism. *Autism, 8*, 349-367.

Sherer, M. R., & Schreibman, L. (2005). Individual behavioral profiles and predictors of treatment effectiveness for children with autism. *Journal of Consulting and Clinical Psychology, 73*, 525-538.

Smith, T., Scahill, L., Dawson, G., Guthrie, D., Lord, C., Odom, S., Rogers, S., & Wagner, A. (2007). Designing research studies on psychosocial interventions in autism. *Journal of Autism and Developmental Disorders, 37*, 354-366.

(6) Whalen, C., & Schreibman, L. (2003). Joint attention training for children with autism using behavior modification procedures. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 44*, 456-468.

White, S. W., Keonig, K., & Scahill L. (2007). Social skills development in children with autism spectrum disorder: A review of the intervention research. *Journal of Autism and Developmental Disorders, 37*, 1858-1868

Wing, L., & Gould, J. (1979). Severe impairments of social interaction and associated abnormalities in children: Epidemiology and classification. *Journal of Autism and Developmental Disorders, 9*, 11-29.

CHAPTER III

EXAMINING THE EFFECTIVENESS OF PEER-MEDIATED AND VIDEO-MODELING SOCIAL SKILLS INTERVENTIONS FOR CHILDREN WITH AUTISM SPECTRUM DISRODER: A META-ANALYSIS IN SINGLE-CASE RESEARCH USING HLM

Introduction

The impairment of social interaction is recognized as one of the core deficits for children with autism spectrum disorders (ASD) (White, Keonig, & Scahill, 2007). The advanced studies in children with ASD have provided a clearer definition for diagnosis and helped clinicians identify children with ASD earlier and more accurately. As a result, more cases with the core deficit in social interaction are being identified, such as children with high function autism (HFA), pervasive developmental disorders not otherwise specified (PDDNOS), or Asperger's disorder. Children with HFA, PDDNOS or Asperger's disorder show fewer symptoms of cognitive or language deficits, but social interaction is a major barrier for them (Rao, Beidel, & Murray, 2008). The practice of inclusive education with its focus of placing children with disabilities in regular classrooms has increased opportunities for ASD children to socialize with peers. However, children with ASD often do not interact effectively or spontaneously with normally developing classmates (Laushey & Heflin, 2000). As a result, social skill intervention can be critical for successful inclusion of children with ASD.

A variety of intervention models have been developed for the social skill training of children with ASD including behavior modification, peer-mediated training, social story, video-modeling, pivotal response training, joint attention training, and buddy system (Bass & Mulick, 2007; Matson, Matson, & Rivet,

2007; Nikopoulos & Keenan, 2004; Scattone, 2007). There is a growing demand for empirical evidence of the effectiveness of those approaches. Parents want to know how to choose an effective model for their children with ASD, clinicians want to adopt effective interventions for their evidence-based practice, and policy-makers are interested in funding programs with proven track record. Furthermore, the No Child Left Behind Act (U.S. Department of Education, 2002) has a strong emphasis on the adoption of scientifically based research findings to guide the instruction and education of the children.

Several studies have evaluated the effectiveness of individual intervention models. The synthesis of the outcomes from these studies would be beneficial for those who are interested in social skills interventions for children with ASD. In addition to the effectiveness of varied approaches, examination of the factors that may impact the effectiveness of the interventions is also important. Kasari, Freeman, et al. (2005) and National Research Council [NRC] (2001) argue that there is no one approach that fits the needs of all children with ASD. Children's progress can be influenced by the factors such as age, mental ability, and language ability as well as the duration and frequency of the intervention (Kasari, Freeman, et al., 2005; Kasari & Rotheram-Fuller, 2007; NRC, 2001). Hence, the synthesis of the intervention studies with the inclusion of the analysis of the factors that possibly moderate intervention outcomes can help to identify the interventions that are most likely to benefit a specific child.

Meta-analysis is recognized as a quantitative and objective tool for assessing the magnitude of the intervention effects through the synthesis of a set of the

studies. According to Glass (1976, p.3), meta-analysis uses diverse statistical methods to synthesize the quantitative outcomes of many independent studies. An effect size (ES) is most often used in meta-analyses as an index of treatment success. This method, however, is applied more often to studies with group-comparison designs and rarely to studies with single-case design (Van den Noortgate & Onghena, 2003a & 2003b; Miller, 2005). Intervention studies for children with ASD rarely use a group-comparison design. Matson et al. (2007) reviewed social skill treatment studies in children with ASD and calculated that more than 90% of the reviewed papers adopted single-case designs. The method of single-case research in which the behavior of one or more participants is measured repeatedly over time is preferred in the studies of children with ASD for several reasons. For instance, the relatively low prevalence of ASD makes random assignment of participants to the groups difficult. In addition, assigning children with ASD to the control group may violate their needs for early intervention and raise ethical concerns. More practical reasons why clinicians are attracted to single-case research designs include the relatively low cost of single-case studies and the resulting in-depth information of participant responses (Van den Noortgate & Onghena, 2007).

The assessment of treatment effectiveness in single-case studies typically relies on visual analysis of the outcome graphs (Brossart, Parker, Olson, & Mahadevan, 2006; Van den Noortgate & Onghena, 2003a; Miller, 2005). Visual analysis of the graphs is meant to visually inspect the variability, trend, and overlap of the data points between or within the intervention phases (Miller, 2005).

This method, however, is considered subjective and may be invalid for the comparisons of multiple studies. Low inter-rater agreement has been reported between experts and well-trained raters while they visually examined the effectiveness of multiple studies (Brossart et al., 2006). As a result, meta-analysis, a quantitative synthesis using statistical methods, has been suggested as a preferable alternative as it can result in more objective and comprehensive evaluation of multiple studies (Van den Noortgate & Onghena, 2003a).

There are, however, major challenges with quantification and synthesis of single-case studies. The challenges result from the specific features of the single-case research designs such as different scales used in different studies and the autocorrelation and trend in the data (Jenson, Clark, Kircher, & Kristjansson, 2007; Miller, 2005). In order to integrate data across studies, the raw data from the dependent variables (DVs) from different studies, originally measured on different scales, must first be standardized for further comparison (Raudenbush & Bryk, 2002; Van Den Noortgate & Onghena, 2003b). Several traditional statistical methods have been proposed to provide standardized scales, including percentage of nonoverlapping data points (Scruggs, Mastropieri, & Casto, 1987), effect size calculated by dividing the mean difference between pre-treatment and post-treatment data-points with the standard deviation of pre-treatment data-points (Busk & Serlin,1992), percentage of zero data (Scotti, Evans, Meyer, & Walker, 1991), and different regression methods (e.g., Allison & Gorman, 1993; Faith, Allison, & Gorman, 1996). Although these methods contribute to the standardization and quantification of the intervention outcomes, many of them

have been criticized due to their limitations (Campbell, 2004; Jenson et al, 2007; Miller, 2005). For instance, percentage of zero data cannot be used in calculating the effectiveness of learning behavior because the learning behavior is expected to increase instead of diminishing to zero at treatment phase. The effectiveness of the studies with large variation in targeted behavior such as social performance tends to be underestimated by using percentage of nonoverlapping data. Busk and Serlin's method of calculating the effect size and the regression-based methods have been criticized generally for the difficulties in the interpretation of the effect sizes (Brossart et al., 2006; Campbell, 2004; Miller, 2005).

In addition to the problem with different scales, the second challenge with synthesis of single-case research studies is autocorrelation of data measured at different time points. Autocorrelation is the correlation between the sequential data points that is always assumed to exist in repeated measurements of the same individual over time due to serial dependency of the data. When autocorrelation exists, the assumption of data independence is violated, which may lead to Type I errors (Jenson et al., 2007; Miller, 2005; Raudenbush & Bryk, 2002). Similarly, trend of the data is the upward or downward direction of the data point line regardless of intervention which can also result in the bias of effect size and Type I errors (Miller, 2005).

In order to solve problems associated with the data of single-case studies, Van Den Noortgate and Onghena (2003b, 2007), Raudenbush and Bryk (2002), and Jenson et al. (2007) have suggested the use of Hierarchical Linear Modeling (HLM) to examine the outcomes across multiple single-case studies. Prior to the

HLM analysis, the data from multiple studies are transformed to a standardized metric, so that the outcomes can be compared on the same scale. Furthermore, the scores of HLM are generally structured in units. For instance, the data points from the same DV and the effect sizes from the same participant or the same study in units are coded and analyzed as the same category in HLM. By setting hierarchical structure to the data, the researcher can take the dependence of the scores into consideration that is caused by the influence of being at the same unit or category (Miller, 2005; Van Den Noortgate & Onghena, 2003b & 2007). With HLM, the researchers can manage the scaling and dependence problems of the scores, and also describe the variance at the same level of these units by adopting the characteristics of the units as the predictors. The application of the HLM to the meta-analysis can help estimate the mean and variation of effect size parameters across multiple studies and test how well different predictors can explain the variation (Miller, 2005; Raudenbush & Bryk, 2002; Van Den Noortgate & Onghena, 2007).

This study used HLM to examine the effect sizes of peer-mediated and video-modeling interventions that adopted single-case research design to improve the social behavior for children with ASD. The peer-mediated approach tends to be the most empirically supported model of social skill intervention for children with ASD (Apple, Billingsley, & Schwartz, 2005; Bass & Mulick, 2007). Training peers to support social skills training is assumed to be a more natural method for children and diminish the problem of limited generalizability in adult-mediated interventions (Strain, Schwartz, & Bovey, 2008). In this approach, the peers are

normally developing children who are taught social interaction strategies such as sharing, helping, prompting, instructing, or praising by the researchers. Once the training is completed, the peers are asked to apply those strategies while interacting with children with ASD in order to promote their social interaction. The video-modeling intervention has also been recognized as an effective method to improve the social-communication skills in children with autism (Apple et al., 2005; Bellini & Akullian, 2007). In this approach, children with ASD are asked to observe a video of proper social interaction and to mimic the social behavior demonstrated in the video. The utilization of video media can provoke children's interest and be a natural way for children to learn skills.

We also investigated whether the effect sizes vary as a function of age, treatment approaches (peer-mediated vs. video-modeling), and their interaction. From the review of the existing literature, we expected that both intervention approaches would be effective. However, we did not know if one of the two approaches would be more effective than the other. We expected that age could impact the effectiveness of the interventions significantly. In terms of the interaction of treatment approach and age, we did not have a specific hypothesis.

Method

Identification and selection of reviewed papers.

We searched the Ovid MEDLINE, PsycINFO, ERIC, Web of Science, TOC Premier databases with the keywords: (1) autism or autistic, (2) social or psychosocial, and (3) therapy or training or intervention or treatment. Only papers published between 1994 and January of 2008 and in English were collected. The

reason for setting limitation to 1994 is because the diagnostic criteria between DSM (Diagnostic and Statistical Manual of Mental Disorders) and ICD (International Statistical Classification of Diseases and Related Health Problems) became more consistent after DSM-IV was published in 1994. The first author reviewed all the initial papers (n = 707) and identified 64 papers that met the following criteria: (1) at least half or more of the participants were children diagnosed with ASD, (2) the DVs of the study were related to social skills and involved interaction with human beings, (3) the focus of the study was to provide an intervention, and (4) the researchers had adopted single-case research design in the study. If the researchers had collected both single-case and group-comparison data, there should be individual information and outcome graphs for the participants in the single-case part of the study.

From the 64 papers, 13 papers that adopted either peer-mediated or video-modeling treatment were selected for further analysis. Peer-mediated intervention was defined as an intervention where the core independent variable of the intervention involved training of the peers with the goal to improve social interaction of the children with ASD. In video-modeling interventions, the core independent variable of the training was having children with ASD watch a video and mimic social performance that was demonstrated in the video. The studies that mixed other treatment models with peer-mediated or video-modeling intervention were excluded given that the treatment effects might be confounded. Studies that alternated peer-mediated or video-modeling intervention with other different intervention models were included only if peer-mediated or

video-modeling intervention was implemented in the first treatment phase such as in Thiemann and Goldstein (2004). In total, there were fourteen studies in the thirteen papers because one of the papers included two studies. Among these studies, nine adopted peer-mediated intervention and five adopted video-modeling intervention. Nine of the thirteen papers were published after the year 2000. A summary of the reviewed studies is provided in Table 3.1.

Coding of the studies.

The participant characteristics and data points of the included thirteen papers were coded. It is a common practice for single-case studies to present the results in the format of figures that display the trend of data measured at different time points. To accurately read the values of the data points from a figure, each figure of the included DVs was scanned and imported into Microsoft Office Visio 2007. If there were more than one DV within the study, only the data from those DVs that represented the social behavior of the children with ASD were coded. However, if the DVs within the study represented redundant information, then only the DV representing an overall outcome of the measured behaviors was chosen in the study. For example, Thiemann and Goldstein (2004) reported the DVs of total initiations and responses and the number of initiations, only the DV of total initiations and responses was coded. If there were multiple baseline and intervention phases such as in a reversal design, only the first baseline and intervention phases were coded. Furthermore, the DVs with less than two data points at either baseline or intervention phase were not coded. The data points used to demonstrate the generalization effect (i.e., to measure the behavior at

different settings in order to examine whether the outcome can be generalized) were not coded. Some DVs were reverse coded such that higher values always indicated higher level of social performance. Following the criteria mentioned above, there were 89 DVs coded from 43 participants including 37 males and 6 females. The age of the participants ranged from four to 15 with a mean age of 6.49 (*SD*=2.38). The average age was 6.07 (*SD*=1.65) for the studies using peer-mediated intervention and 7.36 (*SD*=3.37) for the studies using video-modeling intervention.

Transformation of the data points for HLM and effect sizes.

After coding, the data points of each DV were first transformed into standardized scores. The standardized score of each data point was obtained by subtracting the raw score of the data point from the mean of the data points within the baseline and intervention phases and then dividing by the standard deviation of the data points coming from the baseline and intervention phases combined. A total of 1796 standardized scores from 89 DVs and 43 participants were calculated. In order to fit the data with the proposed HLM model, these standardized scores were further subtracted from the mean of the standardized data points at the corresponding baseline phase in order to ensure the mean of the data points in the baseline phase equal to zero. After this procedure, the effect size of the dependent variable would be equal to the mean of the subtracted scores of the standardized data points at the corresponding intervention phase.

Examining effect sizes with HLM.

The final data file included the standardized scores of the data points, the corresponding (either baseline or intervention) phases, participants ID, participants' age, and the intervention approach (either peer-mediated or video-modeling). All analysis were computed with HLM 6 (Raudenbush, Bryk, Cheong, & Congdon, 2004). The following three-level HLM model was used in this study.

Level-1 Model

$$ZDATAPOI_{ijk} = \pi_{1jk} *(PHASE_{ijk}) + e_{ijk}$$

Level-2 Model

$$\pi_{1jk} = \beta_{10k} + r_{1jk}$$

Level-3 Model

$$\beta_{10k} = \gamma_{100} + \gamma_{101} (TX) + \gamma_{102} (AGE) + \gamma_{103} (TX*AGE) + \mu_{10k}$$

Level-l model represents a regression equation for each DV. In the Level-1 model, the outcome variable, $ZDATAPOI_{ijk}$, is the standardized score of the data point for occasion i, DV j, and participant k; π_{1jk} is the effect size of each DV that equals the difference between the intervention and baseline means of the standardized scores; $PHASE_{ijk}$ is a dichotomous variable that reflects the phase of each data point (i.e., 1 indicates the data point is in the intervention phase and 0 means the data point is in the baseline phase); e_{ijk} is a random error term. Level-2 model indicates the effect size for each DV, π_{1jk}, equals the effect size of its corresponding participant (i.e., β_{10k}) plus a residual (i.e., r_{1jk}). With the effect sizes for DVs from HLM, the effect sizes for each participant can be calculated by averaging the effect sizes of the DVs measured upon the participant. Furthermore,

the effect sizes for each study can be calculated by averaging the effect sizes of the participants in the study. In Level-3 model, β_{10k} is the effect size for each participant. γ_{100} equals the grand mean effect size across all participants and DVs. TX is the coding of treatment type. AGE is the age of the participant. TX*AGE indicates the interaction of treatment type and participant's age. μ_{10k} is an error term. Level-3 model was used to examine the effects of treatment type, participant's age, and the interaction of treatment type and age on the effect sizes.

Results

The effect sizes of the reviewed studies ranged from 0.65 to 2.31with the mean of 1.27 (*SD* = 0.43, 95% CL=1.05~1.50). Twelve of the 14 studies, 37 of the 43 participants, and 67 of the included 89 DVs yielded large effect sizes according to the Cohen's (1988) guidelines. The results of the HLM analysis are displayed in Table 3.2. The mean of the effect sizes ($\gamma_{100)}$ was significantly different from zero suggesting that the examined social skills interventions are effective in improving the social behavior of participating children with ASD. The effect of treatment type was minimal, suggesting that both peer-mediated and video-modeling approaches were equally effective. Additionally, age predicted effect size significantly. The negative coefficient of -0.05 for the age factor indicates that interventions tended to be slightly more effective for younger children. The interaction effect of treatment type and age approached significance (p= 0.066) suggesting that the impact of age may vary depending on the type of the intervention. This interaction is depicted in Figure 1, which shows that both the age effect and the age by treatment interaction effect may have been produced by

the single video-modeling study (Nikopoulos & Keenan, 2003) that included older participants. Peer-mediated studies show no clear age effect, but none included participants older than 10 years-of-age.

Discussion

The results of the current study indicate that peer-mediated and video-modeling interventions are both effective in improving social behavior of children with ASD and there is no significant difference between the effectiveness of these two intervention approaches. Participant's age impacts the intervention outcomes significantly. Our results suggest tentatively that younger child may benefits more from the interventions than the older children. Additionally, participant's age may interact with intervention approaches and impact intervention effectiveness.

The significant grand mean effect size from the HLM analysis suggested that both peer-mediated and video-modeling approaches can significantly improve the social performance for children with ASD. Using Cohen's standards, over 75 % of the effect sizes of the included studies were large. These results are notable given that meta-analyses of social skill interventions have not usually yielded a positive or large effect size. Minimal effectiveness or small effect sizes of social skill interventions have been reported irrespective of the differences in intervention settings, meta-analysis methods, research designs, or targeted population of the studies. For example, Bellini, Peters, Benner, and Hopf (2007) examined 55 single-case studies of social skill interventions in a school setting for subjects with ASD using percentage of non-overlapping data points as the measurement of

effect sizes. They calculated the effect sizes of the studies based on their intervention types (e.g., collateral, peer-mediated, child-specific, or comprehensive), formats (e.g., individual or group), location (e.g., classroom or pullout), and age groups (e.g., preschool, elementary, or secondary) and reported that the effectiveness of interventions was low to questionable across different categories. As well, a small effect size was found in the meta-analysis of group-comparison studies that focused on social skill training for children with emotional and behavior disorder (Quinn, Kavale, Mathur, Rutherford, & Forness, 1999). Similarly, the review of experimental studies in social skill intervention for student with learning disability yielded a small effect size (Kavale & Moster, 2004).

In the comparison of peer-mediated and video-modeling intervention, the difference between the effect size estimates of the two models was minimal, suggesting that the peer-mediated interventions are as effective as the video-modeling interventions. While the number of studies is still relatively small, these results suggest that both approaches can be recommended as evidence-based practices for social skill training of children with ASD.

Age was found to show a significant moderating effect on the effectiveness of the intervention. A negative coefficient of -0.05 for the age variable suggested the effectiveness of the intervention tends to decrease as a child grows older. This finding was consistent with the previous studies that indicated there were differential gains from the intervention, with younger children tending to gain more from the intervention than older children (Baker-Ericzen, Stahmer, & Burns, 2007; Corsello, 2005). The negative relationship between participant's age and

intervention effectiveness supports the importance of early intervention. However, these interventions can still result in positive gains of social skills for older children with ASD. The effect size of the intervention is still larger than 1.00 for the 10-year-old child because the effect size decrease only -0.05 as the child gets one year older.

In addition, the significant interaction effect between intervention type and age implies that the impact of age on intervention outcomes depends on what type of intervention approaches is adopted and the impact of intervention type on intervention outcomes varies depending on participant's age. According to Figure 3-1, age tends to play a more significant role in the video-modeling approach suggesting that the older children progress less than the younger children. However, the impact of age on effect size in the video-modeling approach is supported by a limited number of studies. Especially, the impact of age between 10 to 15 year-old on effect sizes can be only based on the data coming from a single study. Therefore, more studies that apply video-modeling intervention to older children should be included in the future meta-analysis in order to confirm the finding. In contrast, there is no clear negative relationship between age and effect sizes in the peer-mediated intervention. However, the studies of the peer-mediated intervention did not include the participants older than age 10. Therefore, there is no data that support similar gains of the peer-mediated intervention for children older than 10 year of age. Clearly, more studies and extended range of participants' age are needed in order to clarify the interaction effect of participant's age and interventions on the intervention effectiveness.

Special attention should be paid to the effect size interpretation in single case research and the limitations of the present study. Although relative large effect sizes were reported in the present study and that finding was similar to ones in most of the single-case studies, applying Cohen's guidelines to the interpretation of effect sizes from single-case studies may be inappropriate as Cohen's guidelines are developed originally for between-group designs (Beeson & Robey, 2006). For example, the magnitude of effect size in single-case studies is very likely to be larger than 0.8 and the value of the effect size ranging from 1 to 2.85 or even larger is often seen in single-case research. These large effect sizes are not frequently found in quasi-experimental studies. Beeson and Robey (2006) suggested there should be different benchmarks for the effect size of a particular intervention. The gains of the intervention can be significant in the group of children with mild deficits but not in the one with profound impairment. Similarly, the goal targeted on social initiation may not be reached as easily as the goal focused on social response. As a result, developing different standards for the magnitude of effect size according to specific research designs, types of participants and intervention goals is recommended (Beeson & Robey, 2006; Brossart, et al., 2006).

In addition, the concerns with the order effects of withdrawal or multiple interventions limited the number of the data points included for data analysis. The power of the analysis can be limited due to the excluded data points. Moreover, the present study did not examine several predictors such as categories of DVs, intensity of the intervention, experimental settings, or participant's abilities that

may account for variation in effect sizes. Inclusion and analysis of additional predictors and their interaction can help us identify active ingredients that can lead to best outcome of intervention. For example, children with different levels of cognitive function may response to same intervention differently. Certain amount of training hours may be necessary in order to benefit from the intervention.

The inclusion of additional predictors depends on the progress of the research in children with ASD. For instance, the definition of social behavior must be clarified in order to categorize outcome variables. In addition, sufficient information of experimental setting, intervention duration or frequency, and participant's abilities must be available across different studies. Further meta-analyses of the intervention outcomes will benefit from the fully contextualized information of the studies. As a result, setting up the guidelines for the definition of social behavior and provision of the study information is strongly recommended for the research of children with ASD. Furthermore, the effectiveness of the interventions should not only depend on the magnitude of effect sizes. Adoption of comprehensive criteria and procedures that includes the review of the methodological quality of the studies is recommended (Cook, Tankersley, & Landrum, 2009).

This study used HLM to examine the effectiveness of peer-mediated and video-modeling approaches in social skill training of children with ASD. Both approaches were found to be effective. In order to benefit more from the intervention, it is recommended that social skill training, especially the

video-modeling intervention, should be provided to children with ASD as soon as they are identified.

Footnote. A version of this chapter has been published. Wang, Cui, & Parrila 2011. Research in Autism Spectrum Disorder. 5: 562-569.

Table 3.1.

Summary of the reviewed studies

First author	PY	Children with ASD	Males	Females	Age (mean)	Age (min, max)	Research design(s)	Effect size	Intervention model
Apple, A. L. –exp1	2005	2	2	0	5	5	3+2	0.99	Video-modeling
Apple, A. L. –exp2	2005	2 (3) [a]	1 (2) [a]	1	4.5	4-5	3	0.93	Video-modeling
Chung, K.	2007	4	4	0	6.5	6-7	1	1.16	Peer mediated
Gonzalez-Lopez, A.	1997	4	2	2	6	5-7	2	0.71	Peer mediated
Kohler, F. W.	1995	3	3	0	4	4	2	1.48	Peer mediated
Kohler, F. W.	2007	1	0	1	4	4	3	1.63	Peer mediated
Laushey, K. M.	2000	2	2	0	5	5	2	1.64	Peer mediated
Maione, L.	2006	1	1	0	5	5	3	1.22	Video-modeling
Mundschenk, N. A.	1995	3	2	1	8.67	7-10	2	1.47	Peer mediated
Nikopoulos, C. K.	2003	5 (7) [a]	4 (6) [a]	1	11.2	9-15	1,2	0.65	Video-modeling
Pierce, K.	1997	2	2	0	7.5	7-8	3	1.41	Peer mediated
Reeve, S. A.	2007	4	3	1	5.75	5-6	1	2.31	Video-modeling
Strain, P.S.	1995	5	5	0	4.8	4-6	2	1.11	Peer mediated
Thiemann, K.S.	2004	5	5	0	7	6-9	3	1.09	Peer mediated

Note. PY = Publication year; Research design(s): 1. AB; 2. Reversal; 3. Multiple-baseline; 4. Alternating.

[a] the number outside the parenthesis is the number of the children included in the meta-analysis, and the number inside the parenthesis indicates the original number of the children in the study.

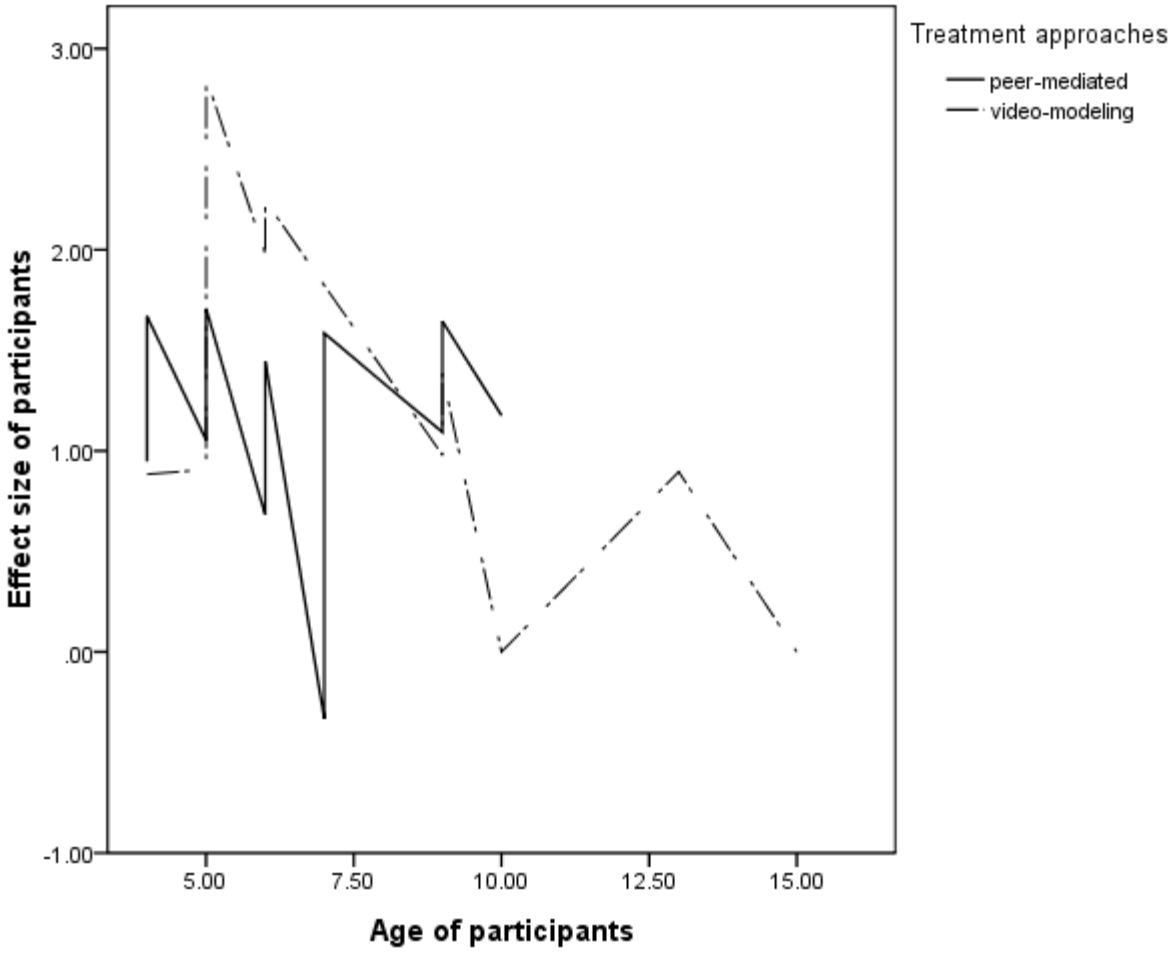
Table 3.2.

Results of 3-level HLM final estimation of fixed effects (with robust standard errors)

Fixed Effect	Coefficient	SE	T-ratio	Approx. *d.f.*	*P*-value
For PHASE slope, π_{1jk}					
For INTRCPT2, β_{10k}					
INTRCPT3, γ_{100}	1.232	0.076	16.035	39	0.000
TREATMENT, γ_{101}	0.002	0.076	0.032	39	0.975
AGE, γ_{102}	-0.051	0.024	-2.137	39	0.039
INTERACT, γ_{103}	0.045	0.024	1.892	39	0.066

Figure 3-1:

The interaction effect of treatment type and age on effect sizes



References

Allison, D. B., & Gorman, B. S. (1993). Calculating effect sizes for meta-analyses: The case of the single case. *Behavior Research and Therapy, 31,* 621-631.

*Apple, A. L., Billingsley, F. & Schwartz, I. S. (2005). Effects of video modeling alone and with self-management on compliment-giving behaviors of children with high-functioning ASD. *Journal of Positive Behavior Interventions, 7,* 33-46.

Baker-Ericzen, M.J., Stahmer, A.C., & Burns A. (2007). Child demographics associated with outcomes in a community-based pivotal response training program. *Journal of Positive Behavior Interventions, 9,* 52 – 60.

Bass, J. D., & Mulick J.A. (2007). Social play skill enhancement of children with autism using peers and siblings as therapists. *Psychology in the School, 44,* 727-734.

Beeson, P. M., & Robey, R. R. (2006). Evaluating single-subject treatment research: Lessons learned from the aphasia literature. *Neuropsychology Review 16,* 161-169.

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,* 264–287.

Bellini, S., Peters, J. K., Benner, L., & Hopf, A. (2007). A meta-analysis of school-based social skills interventions for children with autism spectrum disorders. *Remedial and Special Education, 28,* 153-162.

Brossart, D. F., Parker, R. I., Olson, E. A., & Mahadevan, L. (2006). The relationship between visual analysis and five statistical analyses in a simple AB single-case research design. *Behavior Modification, 30*, 531-563.

Busk, P. L., & Serlin, R C. (1992). Meta-analysis for single-case research. In T. R. Kratochwill & J. R. Levinn (Eds.), *Single-case research design and analysis: New directions for psychology and education* (pp. 187-212). Hillsdale, NJ: Erlbaum.

Campbell, J. M. (2004). Statistical comparison of four effect sizes for single-subject designs. *Behavior Modification, 28*, 234-246.

*Chung, K., Reavis, S. Mosconi, M., Drewry, J., Matthews, J., & Tasse, M. J. (2007). Peer-mediated social skills training program for young children with high-functioning autism. *Research in Developmental Disabilities*. *28*, 423-436.

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

Cook, B. G., Tankersley, M., & Landrum, T. J. (2009). Determining evidence-based practices in special education. *Exceptional Children**,** 75*, 365-383.

Corsello, C. M. (2005). Early intervention in autism. *Infant and Young Children*, *18*, 74-88.

Faith, M. S., Allison, D. B., & Gorman, B. S. (1996). Meta-analysis of single-case research. In R. D. Franklin, D. B. Allison, & B. S. Gorman (Eds.), *Design and analysis of single-case research* (pp. 245-277). Hillsdale, NJ: Erlbaum.

Glass, G. (1976).Primary, secondary, and meta-analysis of research. *Educational Researcher, 5*, 3-8.

*Gonzalez-Lopez, A., & Kamps, E. M.. (1997). Social skills training to increase social interactions between children with autism and their typical peers. *Focus on Autism & Other Developmental Disabilities, 12*, 2-14.

Jenson, W. R., Clark, E., Kircher, J. C., & Kristjansson, S. D. (2007) Statistic reform: evidence-based practice, meta-analyses, and single subject designs. *Psychology in Schools, 44*, 483-493.

Kasari, C., Freeman, S., Paparella, T., Wong, C., Kwon, S., & Gulsrud, A. (2005). Early intervention in autism: Focus on core deficits. *Clinical Neuropsychiatry, 2*, 380-388.

Kasari, C. & Rotheram-Fuller, E. (2007). Peer relationships of children with autism: Challenges and interventions. In E. Hollander & E. Anagnostou (Eds.) *Clinical manual for the treatment of autism*. Arlignton, VA: American Psychiatric Publishing, Inc..

Kavale, K. A., & Moster, M. P. (2004). Social skills interventions for individuals with learning disabilities. *Learning Disability Quarterly, 27*, 31-43.

*Kohler, F. W., Greteman, C., Raschke, D., & Highnam, C. (2007). Using a buddy skills package to increase the social interactions between a preschooler with autism and her peers. *Topics in Early Childhood Special Education, 27*, 155-163.

*Kohler, K. W., Strain. P. S., Hoyson, M., Davis, L., Donina, W. M., & Rapp, N. (1995). Using a group-oriented contingency to increase social interactions

between children with autism and their peers- a preliminary-analysis of corollary supportive behaviors. *Behavior Modification, 19,* 10-32.

*Laushey, K. M., & Heflin, L. J. (2000). Enhancing social skills of kindergarten children with autism through the training of multiple peers as tutors. *Journal of Autism and Developmental Disorders, 30,* 183-193.

*Maione, L., & Mirenda, P. (2006). Effects of video modeling and video feedback on peer-directed social language skills of a child with autism. *Journal of Positive Behavior Interventions, 8,* 106-118.

Matson, J. L., Matson, M. L., & Rivet, T. T. (2007). Social-skills treatments for children with autism spectrum disorders an overview. *Behavior Modification, 31,* 682-707.

Miller, A.L. (2005). *Interventions targeting reciprocal social interaction in children and young adults with autism spectrum disorders: A meta-analysis.* Unpublished dissertation, University of Utah.

*Mundschenk, N. A., & Sasso. G.S.. (1995). Assessing sufficient social exemplars for students with autism. *Behavioral Disorders, 21,* 62-78.

National Research Council (NRC). (2001). *Educating children with autism.* Washington, DC: National Academy Press.

*Nikopoulos, C. K., & Keenan, M. (2003). Promoting social initiation in children with autism using video modeling. *Behavioral Interventions, 18,* 87-108.

Nikopoulos, C. K., & Keenan, M. (2004). Effects of video modeling on social initiations by children with autism. *Journal of Applied Behavior Analysis, 37,* 93-96.

*Pierce, K., & Schreibman, L. (1997). Multiple peer use of pivotal response training to increase social behaviors of classmates with autism: Results from trained and untrained peers. *Journal of Applied Behavior Analysis, 30*, 157-160.

Rao, P. A., Beidel, D. C., & Murray, M. J. (2008). Social skills interventions for children with Asperger's syndrome or high-functioning autism: a review and recommendations. *Journal of Autism and Developmental Disorder*, *38*, 353-361.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical Linear Models* (2nd ed.). Newbury Park, CA: Sage Publications.

Raudenbush, S. W., Bryk, A. S., Cheong, Y. F., & Congdon Jr., R. T. (2004). *HLM 6: Hierarchical linear and nonlinear modeling*. Lincolnwood, IL: Scientific Software International.

*Reeve, S. A., Reeve, K. F., Townsend, D. B., & Poulson, C. L. (2007). Establishing a generalized repertoire of helping behavior in children with autism. *Journal of Applied Behavior Analysis*, *40*, 123-136.

Quinn, M. M., Kavale, K. A., Mathur, S. R., Rutherford, R. B., Jr., & Forness, S. R. (1999). A meta-analysis of social skill interventions for students with emotional or behavioral disorders, *Journal of Emotional and Behavioral Disorders, 7,* 54-64.

Scattone, D. (2007). Social skills interventions for children with autism. *Psychology in the Schools, 44*, 717-726.

Scotti, J. R., Evans, L. M., Meyer, L. H., & Walker, P. (1991). A meta-analysis of intervention research with problem behavior: Treatment validity and standards of practice. *American Journal of Mental Retardation, 96,* 233-256.

Scruggs, T. E., Mastropieri, M. A., & Casto, G. (1987). The quantitative synthesis of single-subject research: Methodology and validation. *Remedial and Special Education*, *8*(2), 24-33.

*Strain, P. S., Danko, C. D., & Kohler, F. (1995). Activity engagement and social-interaction development in young-children with autism- an examination of "free" intervention effects, *Journal of Emotional and Behavioral Disorders, 3*, 108-123.

Strain, P. S., Schwartz, I. S., & Bovey, E. H. (2008). Social competence interventions for young children with autism. In W.H. Brown, S.L. Odom, S.R. McConnell (Eds.), *Social competence of young children: Risk, Disability, & Intervention* (pp. 253-272). Baltimore: Paul H. Brookes.

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

U.S. Department of Education. (2002). No Child Left Behind Act of 2001. Retrieved March 29, 2009, from http://www.ed.gov/policy/elsec/leg/esea02/107-110.pdf.

Van den Noortgate, W., & Onghena, P. (2003a). Combining single-case experimental data using hierarchical linear models. *School psychology quarterly, 18*, 325-346.

Van den Noortgate, W., & Onghena, P. (2003b). Hierarchical linear models for the quantitative integration of effect sizes in single-case research. *Behavior Research Methods, Instrument, & Computer, 35*, 1-10.

Van den Noortgate, W., & Onghena, P. (2007). The aggregation of single-case results using hierarchical linear models. *The Behavior Analyst Today, 8*, 196-208.

White, S. W., Keonig, K. & Scahill L. (2006). Social skills development in children with autism spectrum disorder: a review of the intervention research. *Journal of Autism and Developmental Disorder, 37*, 1858-1868.

CHAPTER Ⅳ

META-ANALYSIS OF SOCIAL SKILLS INTERVENTIONS OF SINGLE-CASE RESEARCH FOR INDIVIDUALS WITH AUTISM SPECTRUM DISORDER: RESULTS FROM THREE-LEVEL HLM

Introduction

The synthesis of social skills interventions (SSIs) has become a critical issue for individuals with autism spectrum disorders (ASD) as there are increasing demands for effective SSIs in this population. Significant impairments in social interaction have been recognized as the key problem for individuals with ASD since the initial case reported by Kanner (1949). Difficulties in social interaction are the critical symptoms shared between the clients of the five subcategories under ASD listed in the International Classification of Diseases, Tenth Revision (ICD-10) (World Health Organization [WHO], 1992) and the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR) (American Psychiatric Association [APA], 2000). In fact, the unique prototype of social impairment in ASD, other than echolalia, mannerism, or unusual interest, is crucial to the differentiation of ASD from other disorders (Carter et al., 2005; Reichow & Volkmar, 2010). Compared with typical autism, these individuals show fewer symptoms of cognitive or language deficits, but social interaction is a major barrier for them (Rao, Beidel, & Murray, 2008). SSIs are expected to help many individuals with ASD by removing or reducing their major social interaction difficulties.

As social interaction issues in individuals with ASD are highlighted, much attention has been lately given to SSIs for them (Reichow & Volkmar, 2010). A

wide range of intervention models have been proposed to improve social skills of individuals with ASD, including behavior modification, peer-mediated training, social story, video-modeling, pivotal response training, joint attention training, and buddy system (e.g., Bass & Mulick, 2007; Matson et al., 2007; Nikopoulos & Keenan, 2004; Scattone, 2007). Subsequently, the quality and efficacy of different models has become a concern (Cotugno, 2009; Krasny, Williams, Provencal, & Ozonoff, 2003) and comparative systematic reviews of these studies are limited (Matson et al., 2007; Krasny et al., 2003). A synthesis of SSI studies promotes evidence-based practice by providing useful information to clinicians for the adoption of effective interventions and to policy-makers aiming to fund empirically-supported programs.

There are two common directions in study synthesis for probing evidence-based interventions: Quality examination and meta-analysis. Quality examination checks the study quality from the perspective of research methodology. Generally, validity of the studies is investigated using various quality indicators. The scores or ranks on the quality indicator checklist are then used to reflect how valid the outcomes may be. For instance, providing sufficient information for replication, controlling confounding factors, and implementing the study reliably to demonstrate the cause-and-effect relationship between independent and dependent variables are listed as important quality indicators for single-case studies (Horner et al., 2005; Reichow, Volkmar, & Cicchetti, 2008; Wang & Parrila, 2008). Reichow et al. (2008) and Wang and Parrila (2008) provided lists of quality indicators to examine either single-case and/or

experimental studies in the field of autism research. They both stressed that results from studies scoring high on quality checklists deserve more attention than results from studies scoring low.

Meta-analyses, in turn, use diverse statistical methods to assess the magnitude of the intervention effects through the synthesis of quantitative outcomes of many independent studies (Glass, 1976). An effect size (ES) is most often used in meta-analyses as an index of treatment success. In addition to examination of overall intervention outcomes, meta-analysis can investigate whether different variables can account for variability in effect sizes (Wang, Cui, & Parrila, 2011). For example, Wang et al. (2011) analyzed SSI single-case studies for individuals with autism and reported that participants' age impacted treatment outcomes. Similarly, Baker-Ericzen, Stahmer, and Burns (2007) identified that younger children and children with less affected social skills at initial stage would improve more at the community-based parent education of pivotal response training. Examination of the relationships between possible predictors and effect sizes can help to identify the critical elements of successful interventions, and may lead to a better match between intervention characteristics and the needs of individual children.

Traditional methods used in quality examinations and meta-analyses have been developed for group-comparison studies that are different from the single-case designs used in most of the SSI studies of individuals with ASD. In a recent review of SSI studies for individuals with ASD, Matson et al. (2007) noted that more than 90% of them adopted single-case designs. Similarly, 85 % of the

SSI studies reviewed by Reichow and Volkmar (2010) adopted single-case designs. Several reasons, such as the relatively low prevalence of ASD, ethical concerns of violating the needs of individuals with ASD for early intervention, and relatively low cost and convenience of implementation, have attracted professionals in the field to adopt single-case research.

Due to uniqueness of single-case research, both the quality indicators and the meta-analysis methods require modifications. In terms of quality indicators, greater demands need to be placed on the provision of in-depth information on participants, settings, materials, independent variables (IV), and dependent variables (DV) in single-case studies because the detailed information is fundamental to interpretation of the results, to external validity, and to replications that are needed to test the generalizability of the results. Additionally, use of multiple-baseline or reversal design is taken as a quality indicator. Those two types of single-case research designs can provide data that supports a functional relationship between the treatment and the outcome behaviors, thus decreasing threats to internal validity and providing a more powerful statement for the intervention efficacy (NRC, 2001; Richards, Taylor, Ramasamy, & Richards, 1999; Smith et al., 2007).

Similarly, traditional meta-analysis methods need to be adjusted to deal with the challenges that single-case studies create. Jenson, Clark, Kircher, and Kristjansson (2007) and Miller (2005) indicated that there are unique problems, such as the different scales used in different studies and the autocorrelation and trend in the data, in quantitative synthesis of single-case studies. Several statistical

methods, including percentage of nonoverlapping data points (Scruggs, Mastropieri, & Casto, 1987), effect sizes calculated by dividing the mean difference between pre-treatment and post-treatment data-points with the standard deviation of pre-treatment data-points (Busk & Serlin, 1992), percentage of zero data (Scotti, Evans, Meyer, & Walker, 1991), and various regression methods (e.g., Allison & Gorman, 1993; Faith, Allison, & Gorman, 1996), have been developed to deal with these issues. However, these methods have been criticized due to their limitations, such as the inability to apply to learning behavior, underestimation of social behavior outcomes, and difficulties in interpreting the effect sizes (e.g., Brossart, Parker, Olson, & Mahadevan, 2006; Campbell, 2004; Miller, 2005).

Van Den Noortgate and Onghena (2003b, 2007), Raudenbush and Bryk (2002), and Jenson et al. (2007) have all suggested using Hierarchical Linear Modeling (HLM) to solve problems associated with the data from single-case studies. The different scaling issue can be solved by transforming the data from multiple studies to a standardized metric prior to HLM analysis. The data in HLM can be structured into levels (i.e., student, classroom, and school) and the characteristics from these levels (i.e., student's age, inclusive education versus special education, and public school versus private school) can be added as predictors for further analysis. Therefore, the researcher can use HLM to examine the moderating effects of these characteristics on study outcomes. By setting hierarchical structure to the data, the researcher can consider the dependence of the scores as well as test how well different predictors can explain the variation (Miller, 2005; Van Den Noortgate & Onghena, 2007).

As appropriate review and meta-analysis tools have been under construction, the synthesis of the quality and effectiveness of SSIs studies for individuals with autism is still sparse. For instance, Baker-Ericzen et al (2007) and Corsello (2005) identified the importance of age factor through their synthesis of intervention programs for individuals with autism. However, Baker-Ericzen et al (2007) did not focus SSI or social skill progress; instead, their analysis was based on children's adaptive progress at the community programs of a specific training approach. Corsello (2005) drew their conclusion based on the descriptive review of varied early intervention programs for children with autism. Bellini, Peters, Benner, and Hopf (2007) adopted the statistical method, percentage of nonoverlapping data points, to examine 55 single-case studies offering SSIs for students with ASD in school settings, and reported that the intervention outcomes were less impressive. Reichow and Volkmar (2010) provided more comprehensive synthesis of SSIs for individuals with autism. They first used their protocol of quality examination to screen for quality experimental and single-case SSIs studies, and adopted ES for experimental studies and visual analysis for single-case studies to conclude that two types of SSIs, social skills groups and video modeling, are evidence-based. Finally, Wang et al. (2011) compared effectiveness of peer-mediated (9 studies) and video-modeling (5 studies) approaches and concluded that both approaches seem to improve the social performance of children with ASD. No significant differences were found between the approaches.

In the current study, we used HLM to examine the treatment outcomes of SSI

studies for individuals with ASD that adopted single-case research design as well as to examine the effect of the predictors on intervention effectiveness. The predictors examined included the length of the intervention, the age and gender of the participant, and five quality indicators (see below for details). On the basis of recent more limited studies, we expected the overall effect size to be positive (see e.g., Baker-Ericzen et al., 2007; Corsello, 2005; Wang et al., 2011) and moderated by the age of the participants (e.g., Baker-Ericzen et al., 2007; Wang et al., 2011). In contrast, we did not form specific hypotheses about how the effect sizes would are associated with the intervention length, the gender of participants, and the quality indicators as these potential predictors of intervention effectiveness have not been sufficiently examined in previous studies.

Method

Identification and selection of reviewed papers.

We searched Ovid MEDLINE, PsycINFO, ERIC, Web of Science, and TOC Premier databases to identify potential papers for review using the keywords: (1) autism or autistic, (2) social or psychosocial, and (3) therapy or training or intervention or treatment. We focused on the papers published between 1994 and February of 2012 and in English. We selected studies published after 1993 because there were consistent diagnostic criteria between DSM (Diagnostic and Statistical Manual of Mental Disorders) and ICD (International Statistical Classification of Diseases and Related Health Problems) after DSM-IV was published in 1994. The papers identified using these keywords were further examined against the following additional criteria: (1) at least half or more of the

participants were children diagnosed with ASD, (2) the DVs of the study were related to social skills and involved interaction with human beings, (3) the focus of the study was to provide an intervention, and (4) the researchers had adopted single-case research design in the study. Additionally, we only included the papers that provided individual level information and identifiable outcome graphs for the participants. A total of 113 papers with sufficient data to calculate effect sizes met screening criteria mentioned above. Because two of the 113 papers included two studies, a total of 115 studies were reviewed (see Table 4.1. for details). .

Quality coding.

In order to examine the quality of the studies, five quality indicators were recruited from the checklist developed by Wang and Parrila (2008) after reviewing recent reports done by Horner et al. (2005), Lord et al. (2005), Reichow et al. (2008), and Smith et al. (2007). Table 4.2 shows the quality indicators. The first four quality indicators (QI1 to QI4), with a focus on providing sufficient, in-depth, and replicable information of participants, settings/materials, IV, and DV, were chosen because they are fundamental to both internal and external validity of single-case research. Item Five (QI5), multiple baseline or reversal design, was chosen because those two designs are more likely to supply the necessary data to establish a functional relationship between the intervention and the outcome behavior than simple AB /baseline-intervention design (NRC, 2001; Richards et al., 1999; Smith et al., 2007). For each indicator, a study that met the criteria earned one point and the study that did not meet that criteria received zero points. Therefore, the total scores for the studies ranged from zero to five with the lower

scores corresponding to lower validity. Table 4.1 includes the total quality score for each study; however, the HLM analyses used each score separately rather than the sum (see below for details).

Effect size calculations.

In terms of effectiveness examination, each figure of the included DVs was scanned and imported into Microsoft Office Visio 2007 in order to accurately read the values of the data points from a figure. When there were more than one DV within the study, we only coded the data from those DVs that represented social behavior of the participants with ASD and the DV representing an overall outcome of the measured behaviors. For example, Thiemann and Goldstein (2004) reported the DVs of total initiations and responses and the number of initiations; only the DV of total initiations and responses was coded. In order to minimize the confounding and order effects, the first baseline and intervention phases were coded if there were multiple baselines and intervention phases. However, there were a few exceptions. We coded data points from the last intervention phase instead of the first if the last intervention was the ultimate goal of the study (e.g., Strain, Kohler, Storey, & Danko, 1994) or the first intervention did not apply to all participants (e.g., Kohler, Anthony, Stighner, & Hoyson, 2001). The data points used to demonstrate the generalization effect (i.e., to measure the behavior at different settings in order to examine whether the outcome can be generalized) were not included in this analysis. The only exception was Thorp, Stahmer, and Schreibman's (1995) study because there were different data points for different types of generalization and only two or less data points within each phase for each

generalization; all data points listed in their baseline and intervention phases were used for effect size calculation. Following the criteria mentioned above, 18226 data points were coded for 839 DVs from 343 participants (including 285 males and 54 females). The gender information was missing for four participants and age was not reported for 14 of them. The age of the remaining participants ranged from 9 months to 32 years with a mean of 6.51 years (*SD*=4.23).

Transformation of the data points.

The data points of each DV were transformed into standardized scores after coding. The standardized score was obtained by subtracting the raw score of the data point from the mean of the data points within the baseline and intervention phases combined and then dividing by the standard deviation of these data points. In order to fit the data with the proposed HLM model, these standardized scores were further subtracted from the mean of the standardized data points at the corresponding baseline phase in order to ensure the mean of the data points in the baseline phase equaled zero. The scores of some DVs were further reversed in order to have higher values always indicating higher level of social performance. Using this procedure, the effect size of the dependent variable is equal to the mean of the subtracted scores of the standardized data points at the corresponding intervention phase. A few of the DVs (28 out of 839) yielded negative effect sizes.

Statistical analyses.

All statistical analyses were computed with HLM 6 (Raudenbush, Bryk, Cheong, & Congdon, 2004). First, the data file that included the standardized scores of the data points and the coding of the corresponding (either baseline or

intervention) phases, DVs, participants, and studies were entered into a two-level HLM to examine the effect sizes of all DVs (see the following):

Level-1 Model

$$ZDATAPOI_{ijk} = \pi_{1jk} *(PHASE_{ijk}) + e_{ijk}$$

Level-2 Model

$$\pi_{1jk} = \beta_{10k} + r_{1jk}$$

Level-l model represents a regression equation for each DV. In the Level-1 model, the outcome variable, $ZDATAPOI_{ijk}$, is the standardized score of the data point for occasion i, DV j, and participant k; π_{1jk} is the effect size of each DV that equals the difference between the intervention and baseline means of the standardized scores; $PHASE_{ijk}$ is a dichotomous variable that reflects the phase of each data point (i.e., 1 indicates the data point is in the intervention phase and 0 means the data point is in the baseline phase); e_{ijk} is a random error term. Level-2 model indicates the effect size for each DV, π_{1jk}, equals the effect size of its corresponding participant (i.e., β_{10k}) plus a residual (i.e., r_{1jk}). Using the effect sizes for DVs created from two-level HLM, the effect sizes for each participant can be calculated by averaging the effect sizes of the DVs measured upon the participant. Furthermore, the effect sizes for each study can be calculated by averaging the effect sizes of the participants in the study.

A fully unconditional three-level HLM was proposed then (see the following). The three levels represented the hierarchical structure of DVs (level 1), participants (level 2), and studies (level 3). Furthermore, the characteristics of the studies can be added to the third level later on for further analysis.

Level-1 Model: $Y_{ijk} = \pi_{0jk} + e_{ijk}$

Level-2 Model: $\pi_{0jk} = \beta_{00k} + r_{0jk}$

Level-3 Model: $\beta_{00k} = \gamma_{000} + \mu_{00k}$

Level-l model represents a regression equation for each DV. In the Level-1 model, the outcome variable, Y_{ijk}, is the effect size of DV i, participant j, and study k; π_{0jk} is the mean effect size of participant j in study k, and e_{ijk} is a random error term. With the effect sizes for DVs from HLM, the effect sizes for each participant were calculated by averaging the effect sizes of the DVs measured upon the participant that had been created from two-level HLM. Level-2 model indicates the effect size for each participant, π_{0jk}, equals the effect size of its corresponding study (i.e., β_{00k}) plus a residual (i.e., r_{0jk}). Furthermore, the effect sizes for each study were calculated by averaging the effect sizes of the participants in the study. In Level-3 model, β_{00k} is the effect size for each study. γ_{000} equals the grand mean effect size across all studies, participants, and DVs, and μ_{00k} is an error term.

After analyzing the effect sizes of all studies, participants, and DVs using the above fully unconditional model, several characteristics associated with the DV, participant, and study level were added to this three-level model as predictors respectively in order to investigate their impacts on treatment effectiveness. These characteristics were set to be fixed effects because that can minimize the iterations and increase the reliability in the analysis.

First of all, the length of intervention, defined as the number of intervention sessions, was entered into level one. Therefore, the model became the following.

Level-1 Model: $Y_{ijk} = \pi_{0jk} + \pi_{1jk}(Length) + e_{ijk}$

Level-2 Model: $\pi_{0jk} = \beta_{00k} + r_{0jk}, \pi_{1jk} = \beta_{10k}$

Level-3 Model: $\beta_{00k} = \gamma_{000} + \mu_{00k}, \quad \beta_{10k} = \gamma_{100}$

The predictors of participant's gender and age were added to level two to analyze the effects of participant characteristics that included participants' gender and age. The HLM model became the following.

Level-1 Model: $Y_{ijk} = \pi_{0jk} + e_{ijk}$

Level-2 Model: $\pi_{0jk} = \beta_{00k} + \beta_{01k}(Age) + \beta_{02k}(Gender) + r_{0jk}$

Level-3 Model: $\beta_{00k} = \gamma_{000} + \mu_{00k}, \quad \beta_{01k} = \gamma_{010}, \quad \beta_{02k} = \gamma_{020}$

Furthermore, the predictors of participant's gender and age were examined respectively at level two in order to investigate the variation of effect sizes among participants. The HLM model became the following.

Level-1 Model: $Y_{ijk} = \pi_{0jk} + e_{ijk}$

Level-2 Model: $\pi_{0jk} = \beta_{00k} + \beta_{01k}(Age \text{ or } Gender) + r_{0jk}$

Level-3 Model: $\beta_{00k} = \gamma_{000} + \mu_{00k}, \quad \beta_{01k} = \gamma_{010}$

The characteristics of the studies, represented with the corresponding scores of the five quality indicators (QI1 to QI5) that are listed in Table 4.2, were used in level three. Quality indicators are examined in order to probe the relationship between study quality and intervention effectiveness. Intervention type was not examined in this study because of the difficulties of clearly categorizing the intervention type of all the studies. This was due mainly to two factors: lack of clear definitions for specific intervention methods, and the overlapping of similar

intervention strategies across intervention programs. The model for examining the impact of study quality on effect sizes was listed as the following.

Level-1 Model: $Y_{ijk} = \pi_{0jk} + e_{ijk}$

Level-2 Model: $\pi_{0jk} = \beta_{00k} + r_{0jk}$

Level-3 Model: $\beta_{00k} = \gamma_{000} + \gamma_{001}QI_{1k} + \gamma_{002}QI_{2k} + \gamma_{003}QI_{3k} + \gamma_{004}QI_{4k} + \gamma_{005}QI_{5k} + \mu_{00k}$

Results

Quality indictors.

The results for the five quality indicators are listed in Table 4.1. Only 23.5 % of the studies met all five criteria. However, 72.2 % scored four out of five points; the criterion that 60.9 % of the studies failed was the provision of sufficient information about the participants. On the positive side, 89.6 % of the studies adopted either multiple baseline or reversal design, and 79 to 89 % provided sufficient information about the settings/materials, independent variables, and dependent variables to allow replication.

Effect sizes.

The effect sizes of the 839 dependent variables collected from all studies ranged from -1.71 to 3.56 with the mean of 1.36 (*SD* = 0.68, 95% CL = 1.31-1.40). The effect sizes of the reviewed studies (see Table 4.1) ranged from -0.17 to 2.31with the mean of 1.40 (*SD* = 0.43, 95% CL = 1.32-1.48, *N* = 115). The results of 2-level HLM indicated that there was significant variability between the effect sizes of all included DVs. In terms of participant level, the effect sizes ranged from -0.39 to 2.85 with the mean of 1.37 (*SD* = 0.55, 95% CL = 1.31-1.43, *N* =

343). Overall, 103 of the 115 studies, 294 of the 343 participants, and 689 of the 839 dependent variables showed effect sizes larger than 0.80 that is traditionally considered as large following Cohen's (1988) guidelines.

Table 4.3 shows the results from the fully unconditional model. The mean of the effect sizes (γ_{000}) was significantly different from zero, suggesting that the examined social skills interventions on average were effective in improving the social behavior of these participants with ASD. The χ^2 statistic accompanying these variance components indicated significant variability among participants and studies in the effect sizes.

Table 4.4 shows the results from the HLM model that included the length (defined as the number of intervention sessions) as a Level-1 predictor. The effect sizes did not vary as a function of the intervention length. After length was added to the model, significant variability in effect sizes remained both between studies and between participants.

Table 4.5 shows the results from a model that included the participants' gender and age as Level-2 predictors. The results indicated that neither gender nor age accounted for unique variance in effect sizes; however, the variation of effect sizes among participants was no longer significant.

The results from the HLM model that included quality indicators as Level-3 predictors are shown in Table 4.6. Research design was the only quality indicator that accounted for unique between-study variance in effect sizes; the impact of the other four quality indicators was minimal. The studies that adopted either multiple baseline or reversal design had, on average, better intervention outcomes than the

studies that used other designs. In terms of random effects, there was still significant between-studies variation in effect sizes after adding quality indicators into the model.

Discussion

The results of the current study show that, on average, social skills interventions for individuals with ASD are effective. In addition, our results indicate that studies adopting multiple baseline or reversal design tend to show better outcomes than studies using other single-case research designs. However, the intervention outcomes did not vary as a function of the length of intervention, the gender or age of the participants, or the study quality related to provision of sufficient, in-depth, and replicable information of participants, settings/materials, IV, and DV.

The significant grand mean effect size from the HLM analysis suggests that SSIs used in the included single-case studies can significantly improve the social performance for individuals with ASD. Close to 90 % of the examined effect sizes of the included studies were larger than 0.8 and more than 50 % were larger than 1.46. Our findings are consistent with Reichow and Volkmar's (2010) report, a recent review of SSI studies for people with ASD that also supported the SSI benefits in this population using different synthesis methods. In contrast to our study that analyzed effect sizes and moderators across all reviewed single-case studies, Reichow and Volkmar (2010) screened the studies of varied research designs using quality criteria, and then illustrated the intervention effectiveness using effect sizes for experimental studies and visual analysis, such as the number

of the participants who made progress, for single-case studies.

Significant improvements have not always been reported in reviews of SSI studies. For example, Bellini et al. (2007) reported the intervention effectiveness was low to questionable in their meta-analysis of 55 single-case studies of school-based SSIs for students with ASD. However, the method they adopted, percentage of nonoverlapping data points, has limitations (see e.g., Allison & Gorman, 1993) and may not be appropriate for analyzing learning behavior such as social skills. Similarly, small effect sizes were found in a meta-analysis of group-comparison SSI studies for children with emotional and behavior disorders (Quinn, Kavale, Mathur, Rutherford, & Forness, 1999), as well as in a meta-analysis of experimental studies of SSIs for students with learning disabilities (Kavale & Moster, 2004). Our results indicate that SSIs may be particularly beneficial for children with ASD, and that the beneficial effects can be statistically summarized also across single-case studies.

In this meta-analysis, several predictors including intervention length, age, and gender of the participants, and quality indicators related to provision of sufficient, in-depth, and replicable information of participants, settings/materials, IV, and DV (see Table 4.2) were not found to play significant roles in the intervention effectiveness. Research design was the only predictor in this meta-analysis found to significantly mediate the intervention effectiveness. The studies that adopted either multiple baseline or reversal designs tended to yield better outcomes. The reasons for the association between these two designs and intervention outcomes are unknown. Multiple baseline design or reversal design

have been suggested as more powerful designs to demonstrate the treatment efficacy (NRC, 2001; Smith et al., 2007; Wang & Parrila, 2008). Perhaps the studies that adopted multiple baseline or reversal design may stress research quality more and are more likely to implement the intervention more thoroughly. Given that 90% of the included studies used one of these designs and significant effects of these designs on SSI outcomes were found, further investigation is necessary to clarify how research design and intervention outcomes are linked. Additionally, Table 4.6 indicates that different quality indicators may work in different directions in terms of the intervention outcomes. The coefficients for these factors included both positive and negative values. Therefore, it can be inappropriate to investigate the relationship between study quality and effectiveness using the total scores of all quality indicators.

The increasing intervention length, defined as the number of the treatment sessions, was not found to predict treatment effectiveness. Although other factors such as natural growth may confound the intervention gains found in the studies with more intervention sessions, it is reasonable to assume that the effect of SSI programs increases with intervention length as the participants have more opportunities to practice their social behavior when they receive more intervention sessions. Our results did not support this assumption. However, the interpretation of the impact of intervention length may be limited as our meta-analyses only coded the data-points at the first intervention phase in order to manage the confounding and order effects when multiple interventions were used.

In terms of age effects, a non-significant coefficient for the age variable

suggests that similar gains were found across different ages. This result is in contrast with the several recent studies that have suggested age effects (Baker-Ericzen et al., 2007; Corsello, 2005; Wang et al., 2011). These studies suggested that early identification and intervention is important and more cost-effective. Although age was not a significant predictor in this meta-analysis, it is not clear whether the impact of age on treatment outcomes was diminished due to its interaction with intervention approaches (Wang & Parrila, 2008) as great variation of intervention methods were included in this meta-analysis. Lord et al. (2005) stated that different age or developmental level may impact on the learning of specific skills. Similar, some types of social behavior may be best taught at younger age, but other types of social behavior may gain more easily if they are taught at old age after prerequisite skills are built. The potential age by intervention type by target skill interaction requires further study.

The present study has some limitations. Significant unexplained variance in effect sizes remained between the studies indicating that our predictors were not sufficient and researchers need to continue their search for those characteristics across levels that significantly impact treatment outcomes. Identifying important predictors of treatment effects can lead to better matching of children with ASD to different interventions. Kasari, Freeman, Paparella, Wong, Kwon, and Gulsrud (2005), Kasari and Rotheram-Fuller (2007), and NRC (2001) suggested that age, mental ability, and language ability as well as the intervention duration and frequency can affect response to treatment. However, the identification of these predictors relies heavily on research progress in the SSI single-case studies for

individuals with ASD. For instance, the information of intervention duration and frequency and of participant's abilities, including mental or language ability, is often lacking or inconsistently reported across studies, making the analysis of the effects of these factors difficult. Although more than 89% of the studies met the criteria of IV quality indicator by adopting intervention package or providing clear implementation procedure, intervention duration and frequency was not necessarily stated. Therefore, setting more stringent reporting guidelines could provide better information for future meta-analyses.

The analysis of the impact of intervention approaches on treatment outcomes is an important task; however, we were not able to complete the analysis due to several reasons. First of all, lack of operational definitions of social behaviors and unclear descriptions of intervention approaches did not allow the full analysis of the impact that varied intervention approaches may have had on effect sizes. As well, the studies that mixed several intervention approaches or different generalization sessions can confound the comparisons of intervention effects. Furthermore, the meta-analysis method we used, HLM, had limitations in the analysis of comparing more than two categories. However, researchers who are interested in comparing treatment effects of two different intervention approaches may be able assign a smaller number of SSI studies into groups using clear definitions of different intervention approaches and apply HLM to compare their effects (see Wang et al., 2011).

The overall positive outcomes of SSIs for participants with ASD represented in this review can also be misleading due to publication bias. Publication bias has

been noticed in clinical research trials with the tendency of greater publication rates for the studies with significant positive outcomes (Dwan, Altman, Arnaiz, et al., 2008; Easterbrook, Berlin, Gopalan, & Matthews, 1991). Although there were few negative results reported in the recruited studies, negative effects were noticed only in one of the 115 studies, 7 of the 343 participants, or 28 of the 839 DVs. Furthermore, null results were found in zero of all studies, 9 of the 343 participants, or 25 of the 839 DVs. In order to better understand the overall effects of intervention studies, the authors should be encouraged to submit their high-quality studies with no positive outcomes, and the journals need to consider accepting the studies with negative or null results as long as they are of good quality (Thornton & Lee, 2000).

In terms of the benchmarks for the interpretation of effect sizes, we need more studies on comparability of effect sizes obtained through different methods. Applying Cohen's guidelines to single-case studies may be inappropriate (Beeson & Robey, 2006). The results of the present study are consistent with most other meta-analyses of single-case studies in reporting "large" effect sizes according to Cohen's guidelines. More than 85% of the participants and 89% of the studies had effect sizes larger than 0.8. Cohen's guidelines were originally developed for between-group designs and Beeson and Robey (2006) and Brossart et al. (2006) have suggested developing different standards for the effect size interpretation based on the characteristics of research designs, participants, and treatment targets. It is clear that more effort must be paid to the benchmarks that are appropriate for interpretation of effect sizes, in particular when these are obtained from

single-case studies.

Additionally, effect sizes alone may be insufficient indicators of treatment effectiveness. First of all, intervention outcomes cannot be decided by the magnitude of effect sizes alone without taking the context into consideration (Parker & Hagan-Burke, 2007; Parker, Vannest, et al., 2009). Cook, Tankersley, and Landrum (2009) suggested that comprehensive criteria and procedures are needed with the combination of methodological quality examination. Interpretation of the effect sizes using meta-analysis should consider the factors such as study designs, participant characteristics, statistical methods, treatment settings, and target behaviors (Campbell, 2004; Jenson et al., 2007). As well, combining other methods such as visual analysis or clinical significance to examine treatment outcomes can be beneficial as they can provide useful information and help represent the outcomes from a comprehensive perspective.

In conclusion, our results confirmed the overall benefits of SSIs in improving social behavior of individuals with ASD. While our results do not allow us to recommend any specific approach as evidence-based practice, they do provide strong support for making social skills interventions available for children with ASD.

Footnote. A version of this chapter has been published online. Wang, Parrila, & Cui 2012. Journal of Autism and Developmental Disorders. Advance online publication. doi: 10.1007/s10803-012-1726-2

Table 4.1.

Summary of the Reviewed Studies

Authors (PY)	Children with ASD	Males	Females	Age (mean)	Age (min, max)	Quality score	Effect size
Argott et al. (2008)	3	2	1	12	11-14	4	1.31
Apple et al. –exp1 (2005)	2	2	0	5	5	4	0.99
Apple et al. –exp2 (2005)	2 (3)[a]	1 (2)[a]	1	4.5	4-5	4	0.93
Baker et al. (1998)	3	1	2	6.67	5-8	4	1.77
Banda et al. (2010)	2	2	0	6	6	4	1.80
Belchic & Harris (1994)	3	3	0	4.67	4-5	4	1.69
Bernard-Opitz et al. (2001)	8	6	2	6.88	5-8	0	0.46
Betz (2008)	3	a	a	a	4-5	4	1.84
Bock (2007)	1	1	0	12	12	3	2.19
Boutot et al. (2005)	1	1	0	4	4	0	1.58
Buffington et al. (1998)	4	3	1	5	4-6	4	1.99
Buggey et al. (2011)	4	2	2	3.75	3-4	5	1.23
Caballero & Connell (2010)	3	3	0	4.67	4-5	5	1.76
Cannella-Malone et al. (2010)	2	0	2	10	6-14	4	1.36
Cardon & Wilcox et al. (2011)	6	6	0	2.33	2-3	4	1.55
Charlop et al. (2010)	3	3	0	8.67	7-11	5	1.86
Charlop-Christy et al. (2002)	3	3	0	6.67	3-12	4	0.80
Chung et al. (2007)	4	4	0	6.5	6-7	1	1.14
Crozier & Tincani (2007)	4	4	0	3.67	3-5	4	1.19
Deitchman et al. (2010)	3	3	0	6	5-7	4	1.76
Delano & Snell (2006)	3	3	0	7	6-9	4	1.51
Dodd et al. (2008)	2	2	0	10.5	9-12	5	1.26
Dykstra et al. (2012)	3	1	2	3.67	3-4	5	1.11
Feng (2008)	1	1	0	11	11	5	1.71
Ferraioli & Harris (2011)	4	3	1	3.75	3-5	5	1.23
Finnigan & Starr (2010)	1	0	1	3	3	4	1.72
Ganz et al. (2008)	4	4	0	10.25	8-13	4	1.67
Garfinkle & Schwartz (2002)	3	3	0	4	3-5	4	0.50
Gena et al. (1996)	4	3	1	b	11-18	3	2.02
Gena et al. (2005)	3	2	1	4	3-5	3	1.33
Gena (2006)	4	2	2	4	4	2	1.59
Gonzalez-Lopez & Kamps (1997)	4	2	2	6	5-7	3	0.73
Gutman et al. (2010)	2	2	0	15	15	3	1.69
Hagopian et al. (2009)	1	1	0	13	13	3	1.20

(Table 4.1 continues)

(Table 4.1 continued)

Hancock & Kaiser (2002)	4	3	1	3.25	2-4	3	1.47
Hanley-Hochdorfer et al. (2010)	4	3	1	9.5	6-12	4	0.47
Harper et al. (2008)	2	2	0	8.5	8-9	4	1.51
Hughes et al. (2011)	3	1	2	19	16-21	3	2.01
Hwang & Hughs (2000)	3	3	0	2.67	2-3	4	1.46
Ingersoll et al. (2005)	3	3	0	2.33	2-3	3	1.29
Ingersoll & Schreibman (2006)	5	3	2	2.4	2-3	5	0.93
Ingersoll & Gergans (2007)	3	2	1	2.67	2-3	5	1.22
Ingersoll et al. (2007)	5	5	0	3	2-4	5	1.21
Isaksen & Holth (2009)	4	2	2	3.75	3-5	4	1.80
Jahr et al. (2000)	6	5	1	7.83	4-12	4	1.96
Jerome & Sturmey (2008)	2	2	0	32	32	2	1.57
Jones et al. (2006)	4	4	0	2.25	2-3	4	1.36
Jones et al. (2007)	2	2	0	3	3	5	2.05
Jones-study1 (2009)	2	2	0	3.5	3-4	4	1.34
Jones-study1 (2009)	1	1	0	3	3	4	1.25
Jull (2011)	2	2	0	4.5	4-5	5	1.72
Jung et al. (2008)	3	3	0	5.67	5-6	5	1.95
Jurgens et al. (2009)	1	1	0	3	3	4	1.16
Kern & Aldridge (2006)	4	4	0	3.5	3-4	3	1.82
Kleeberger & Mirenda (2010)	1	1	0	4	4	4	1.82
Koegel et al. (2005)	2	c	1[c]	8.5	8-9	4	1.89
Koegel et al. (2009)	3	3	0	3	3	4	1.70
Kohler et al. (1995)	3	3	0	4	4	5	1.48
Kohler et al. (2001)	4	4	0	4	4	2	1.49
Kohler et al. (2007)	1	0	1	4	4	1	1.64
Kravits et al. (2002)	1	0	1	6	6	3	1.70
Krebs et al. (2010)	2	c	c	9.5	9-10	3	1.73
Kuhn et al. (2008)	2	2	0	7.5	7-8	4	1.41
Lacava et al. (2010)	4	4	0	8	7-9	5	0.76
Laushey & Heflin (2000)	2	2	0	5	5	5	1.64
Laughey et al. (2009)	4	4	0	b	b	4	1.76
Leaf (2009)	5	5	0	4.8	4-6	5	1.57
Leaf et al.(2010)	3	3	0	6	5-7	5	1.84
Lee et al. (2002)	3	3	0	13.67	7-27	4	1.16
Lee & Sturmey (2006)	3	3	0	17.33	17-18	4	1.26
Liber et al. (2008)	3	3	0	7.33	6-9	4	1.54
Licciardello et al. (2008)	4	3	1	7.25	6-8	4	1.65
Lofin et al. (2008)	3	3	0	9.67	9-10	2	1.42
MacDonald et al. (2009)	2	2	0	6	5-7	4	1.62

(Table 4.1 continues)

(Table 4.1 continues)

Maione & Mirenda (2006)	1	1	0	5	5	4	1.22
Martins & Harris (2006)	3	3	0	3.67	3-4	5	1.44
Marzullo-Kerth et al. (2011)	4	4	0	7.5	7-8	4	1.77
Matson & Francis (1994)	1	1	0	2	2	4	1.26
McDonald & Hemmes (2003)	1	1	0	18	18	3	1.73
Mitchel et al. (2010)	3	1	2	16.67	15-19	2	1.21
Mundschenk & Sasso (1995)	3	2	1	8.67	7-10	4	1.47
Nikopoulos & Keenan (2003)	7	6	1	11	9-15	4	0.25
Nikopoulos& Keenan (2004)	3	3	0	b	7-9	2	0.67
Norris & Dattilo (1999)	1	0	1	8	8	4	-0.17
Owen-Deschryver et al. (2008)	3	3	0	8	7-10	3	0.81
Ozdemir (2008)	3	3	0	5.67	5-6	4	1.19
Pierce & Schreibman (1997)	2	2	0	7.5	7-8	2	1.42
Randolph et al. (2011)	3	2	1	5.33	3-7	4	0.62
Reagon & Higbee (2009)	3	3	0	3.67	2-6	4	1.44
Reeve et al. (2007)	4	3	1	5.75	5-6	3	2.31
Reichow & Sabornie (2009)	1	1	0	11	11	4	1.70
Sancho et al. (2010)	2	1	1	5	5	4	1.46
Sansosti & Powell-Smith (2006)	3	3	0	10	9-11	3	1.22
Sansosti & Powell-Smith (2008)	3	3	0	7.67	6-9	5	1.39
Scattone (2008)	1	1	0	9	9	5	1.70
Schertz & Odom (2007)	3	3	0	1.33	1-2	5	1.00
Schrandt et al. (2009)	3	2	0	6.33	5-8	4	1.62
Sigafoos et al. (2009)	1	1	0	15	15	4	1.00
Simpson et al. (2004)	4	2	2	5.5	5-6	3	1.94
Stephens (2008)	4	2	2	6.5	5-8	5	0.94
Stevenson et al. (2000)	4	4	0	12.5	10-15	4	1.73
Strain et al. (1994)	3	3	0	4	4-5	4	1.85
Strain & Danko.(1995)	3	3	0	3.67	3-4	3	1.48
Strain et al. (1995)	5	5	0	4.8	4-6	3	1.11
Strain & Kohler (1995)	3	3	0	3.67	3-4	3	1.08
Tayler & Hoch (2008)	3	2	1	5.33	3-8	4	1.51
Tetreault & Lerman (2010)	3	2	1	5.67	4-5	4	0.73

(Table 4.1 continues)

(Table 4.1 continues)

Thiemann & Goldstein (2001)	5	5	0	8.8	6-12	5	1.55
Thiemann & Goldstein (2004)	5	5	0	7	6-9	5	1.09
Thorp et al. (1995)	3	3	0	7.33	5-9	5	1.37
Tsao & Odom (2006)	4	4	0	4.75	3-7	4	1.12
Vismara & Lyons (2007)	3	3	0	2.33	2-3	5	1.12
Vismara & Rogers (2008)	1	1	0	0.75	0.75	4	1.36
Whalen & Schreibman (2003)	4	2[d]	1[d]	4	4	4	1.81
Yang et al. (2003)	4	2	2	8.25	7-9	3	0.52

Note. PY = Publication year; [a]The number outside the parenthesis is the number of the children included in the meta-analysis, and the number inside the parenthesis indicates the original number of the children in the study;[b]There is no information regarding the age of the participants; c[d] There is no information regarding the gender for one of the participant.

Table 4.2.
Quality Indicators for Quality Examination

QI1: ASD Participants

Sufficient information about and across ASD participants including:

- Gender
- Age
- Ability (i.e., IQ, developmental, academic skill, adaptive skills, language…, at least one type of ability tests including its name and score or index)
- Confirmation of ASD diagnosis using instrument using ADOS, ADI-R, CARS, DSM-III or later version, ICD10 or diagnosed by specialists including psychologist, psychiatrist, pediatrician, or neurologist

QI2: Settings/Materials

Detailed information regarding the settings/materials used for intervention:

- Settings: indicate where is the intervention settings (i.e., living room or kitchen at home, classroom or gym at school)
- Materials: if specific material is required (i.e., the story for social story intervention, the script for social interaction or play, instrument for music therapy, computer program for computer-assisted intervention)

QI3:Independent Variables (IVs)

Sufficient information of IVs for replication (i.e., implementing procedure, manual or standardized procedure for implementation)

QI4:Dependent Variables (DVs)

Operationally- defined (observable, measurable) DVs linked clearly to target behavior

Detailed measuring procedure for replication and to generate quantifiable index

QI5:Research Designs

Using multiple baseline or reversal design

Note. ADOS =Autism Diagnostic Observation Schedule, ADI-R=Autism Diagnostic Interview-Revised , CARS=Childhood Autism Rating Scale, DSM – III = Diagnostic and Statistical Manual of Mental Disorders-3rd version , ICD-10 (International Statistical Classification of Diseases and Related Health Problems-10th version).

Table 4.3.

Results from 3-level fully unconditional HLM model (with robust standard errors).

Fixed Effect	Coefficient	SE	*T*-ratio	Approx. *d.f.*
Grand mean study ES , γ_{000}	1.39	0.041	33.92**	114
Random Effect	*SD*	Variance Component	*df*	χ^2
DV, e_{ijk}	0.526	0.277		
Participants, r_{0jk}	0.200	0.040	228	281.49**
Studies, μ_{00k}	0.355	0.126	114	403.68***

Note. ** $p < .01$; ***$p < .001$

Table 4.4.

Results from conditional 3-level HLM model including intervention length as Level 1 predictor (with robust standard errors)

Fixed Effect	Coefficient	SE	*T*-ratio	Approx. *d.f.*
Grand mean study ES , γ_{000}	1.331	0.053	25.09***	114
Overall mean study ES length slope, γ_{100}	0.005	0.003	1.518	837
Random Effect	*S*D	Variance Component	*df*	χ^2
DV, e_{ijk}	0.524	0.275		
Participants, r_{0jk}	0.201	0.040	228	282.95**
Studies, μ_{00k}	0.355	0.126	114	404.38***

Note. ** $p < .01$; ***$p < .001$

Table 4.5.
Results from conditional 3-level HLM model including gender and age as Level 2 predictor (with robust standard errors)

Fixed Effect	Coefficient	SE	*T*-ratio	Approx. *d.f.*
Grand mean study ES , γ_{000}	1.453	0.105	13.778***	109
Overall mean study, age, γ_{010}	-0.012	0.012	-0.964	322
Overall mean gender, γ_{020}	0.009	0.068	0.130	322
Random Effect	*SD*	Variance Component	*df*	χ^2
DV, e_{ijk}	0.525	0.276		
Participants, r_{0jk}	0.179	0.032	213	243.217
Studies, μ_{00k}	0.354	0.125	109	397.034***

Note. ***$p < .001$

Table 4.6.
Results from conditional 3-level HLM model including quality indicators as Level 3 predictor (with robust standard errors)

Fixed Effect	Coefficient	SE	*T*-ratio	Approx. *d.f.*
Grand mean study ES , γ_{000}	1.130	0.202	5.582***	109
Overall mean study ES				
*QI*1, γ_{001}	0.079	0.078	1.009	109
*QI*2, γ_{002}	-0.040	0.104	-0.383	109
*QI*3, γ_{003}	-0.025	0.110	-0.226	109
*QI*4, γ_{004}	-0.067	0.118	-0.565	109
QI5, γ_{005}	0.378	0.167	2.265*	109
Random Effect	*SD*	Variance Component	*df*	χ^2
DV, e_{ijk}	0.526	0.277		
Participants, r_{0jk}	0.199	0.040	228	281.747**
Studies, μ_{00k}	0.341	0.116	109	385.414***

Note. *QI*1 = description of participants with ASD; *QI*2 = description of settings/materials; *QI*3 = description of independent variables; *QI*4 = quality description of dependent variables; *QI*5 = research design.
Note. * $p < .05$; ** $p < .01$; ***$p < .001$

References

*Studies included in the meta-analysis.

Allison, D. B., & Gorman, B. S. (1993). Calculating effect sizes for meta-analyses: The case of the single case. *Behavior Research and Therapy, 31,* 621-631.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text revision). Washington DC: Author.

*Apple, A. L., Billingsley, F. & Schwartz, I. S. (2005). Effects of video modeling alone and with self-management on compliment-giving behaviors of children with high-functioning ASD. *Journal of Positive Behavior Interventions, 7,* 33-46.

*Argott, P., Townsend, D. B., Sturmey, P., & Poulson, C. L. (2008). Increasing the use of empathic statements in the presence of a non-verbal affective stimulus in adolescents with autism. *Research in Autism Spectrum Disorders, 2,* 341-352.

*Baker, M. J., Koegel, R. L., & Koegel, L. K. (1998). Increasing the social behavior of young children with autism using their obsessive behaviors. *Journal of the Association for Persons with Severe Handicaps, 23,* 300-308.

Baker-Ericzen, M. J., Stahmer, A. C., & Burns A. (2007). Child demographics associated with outcomes in a community-based pivotal response training program. *Journal of Positive Behavior Interventions, 9,* 52 – 60.

*Banda, D. R., Hart, S. L., & Liu-Gitz, L. (2010). Impact of training peers and children with autism on social skills during center time activities in inclusive classrooms. *Research in Autism Spectrum Disorders, 4,*

619-625. doi:10.1016/j.rasd.2009.12.005

Bass, J. D., & Mulick, J. A. (2007). Social play skill enhancement of children with autism using peers and siblings as therapists. *Psychology in the School, 44*, 727-734.

Beeson, P. M., & Robey, R. R. (2006). Evaluating single-subject treatment research: Lessons learned from the aphasia literature. *Neuropsychology Review 16*, 161-169.

*Belchic, J. K. & Harris, S. L. (1994). The use of multiple peer exemplars to enhance the generalization of lay skills to the siblings of children wit autism. *Child & Family Behavior Therapy, 16*(2), 1-25.

Bellini, S., Peters, J. K., Benner, L., & Hopf, A. (2007). A meta-analysis of school-based social skills interventions for children with autism spectrum disorders. *Remedial and Special Education, 28*, 153-162.

*Bernard-Opitz, V., Sriram, N., & Nakhoda-Sapuan, S. (2001). Enhancing social problem solving in children with autism and normal children through computer-assisted instruction. *Journal of Autism and Developmental Disorders, 31,* 377-384.

*Betz, A., Higbee, T. S., & Reagon, K. A. (2008). Using joint activity schedules to promote peer engagement in preschoolers with autism. *Journal of Applied Behavior Analysis, 41*, 237-241.

*Bock, M. A. (2007). A social-behavioral learning strategy intervention for a child with Asperger syndrome - Brief report, *Remedial and Special Education, 28*, 258-265.

*Boutot, E. A., Guenther, T., & Crozier, S. (2005). Let's play: Teaching play skills to young children with autism. *Education and Training in*

Developmental Disabilities, 40, 285-292.

Brossart, D. F., Parker, R. I., Olson, E. A., & Mahadevan, L. (2006). The relationship between visual analysis and five statistical analyses in a simple AB single-case research design. *Behavior Modification, 30,* 531-563.

*Buffington, D. M. Krantaz, P.J., McClannanhan, L. E., & Poulson, C. L. (1998). Procedures for teaching appropriate gestural communication skills to children with autism, *Journal of Autism and Developmental Disorders, 28*, 535-544.

*Buggey, T., Hoomes, G., Sherberger, M. E., & Williams, S. (2011). Facilitating social initiations of preschoolers with autism spectrum disorders using video self-modeling. *Focus on Autism and Other Developmental Disabilities, 26*, 25-36.

Busk, P. L., & Serlin, R. C. (1992). Meta-analysis for single-case research. In T. R. Kratochwill & J. R. Levinn (Eds.), *Single-case research design and analysis: New directions for psychology and education* (pp. 187-212). Hillsdale, NJ: Erlbaum.

*Caballero, A., & Connell, J. E. (2010). Evaluation of the effects of social cue cards for preschool age children with autism spectrum disorders (ASD). *Journal of Behavior Assessment and Intervention in Children, 1*, 25-42.

Campbell, J. M. (2004). Statistical comparison of four effect sizes for single-subject designs. *Behavior Modification, 28*, 234-246.

*Cannella-Malone, H. I., Fant, J. L., & Tullis, C. A. (2010). Using the picture exchange communication system to increase the social communication of two individuals with severe developmental disabilities. *Journal of*

Developmental and Physical Disabilities, 22, 149-163. doi:10.1007/s10882-009-9174-4

*Cardon, T. A., & Wilcox, M. J. (2011). Promoting imitation in young children with autism: A comparison of reciprocal imitation training and video modeling. *Journal of Autism and Developmental Disorders, 41*, 654-666. doi:10.1007/s10803-010-1086-8

Carter, A. S., Davis, N. O., Klin, A., & Volkmar, F. R. (2005). Social development in autism. In F. R. Volkmar, R. Paul, A. Klin, & D. A. Cohen (Eds). *Handbook of autism and pervasive developmental disorders* (3rd Edition, pp.312-333). Hoboken, NJ: John Wiley.

*Charlop, M. H., Dennis, B., Carpenter, M. H., & Greenberg, A. L. (2010). Teaching socially expressive behaviors to children with autism through video modeling. *Education and Treatment of Children, 33*, 371-393.

*Charlop-Christy, M. H., Carpenter, M., Le, L., LeBlanc, L. A., & Kellet, K. (2002). Using the picture exchange communication system (PECS) with children with autism: assessment of PECS acquisition, speech, social-communicative behavior, and problem behavior. *Journal of Applied Behavior Analysis, 35*, 213-231.

*Chung, K., Reavis, S., Mosconi, M., Drewry, J., Matthews, J., & Tasse, M. J. (2007). Peer-mediated social skills training program for young children with high-functioning autism. *Research in Developmental Disabilities. 28*, 423-436.

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

Cook, B. G., Tankersley, M., & Landrum, T. J. (2009). Determining

evidence-based practices in special education. *Exceptional Children, 75*, 365-383.

Corsello, C. M. (2005). Early intervention in autism. *Infant and Young Children, 18*, 74-88.

Cotugno, A. J. (2009). Social competence and social skills training and intervention for children with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 39*, 1268-1277.

*Crozier, S., & Tincani, M..(2007). Effects of social stories on prosocial behavior of preschool children with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 37*, 1803-1814.

*Deitchman, C., Reeve, S. A., Reeve, K. F., & Progar, P. R. (2010). Incorporating video feedback into self-management training to promote generalization of social initiations by children with autism. *Education and Treatment of Children, 33*, 475- 488.

*Delano, M., & Snell, M. E. (2006). The effects of social stories on the social engagement of children with autism. *Journal of Positive Behavior Interventions. 8*, 29-42.

*Dodd, S., Hupp, S. D. A., Jewell, J. D., & Krohn, E. (2008). Using parents and siblings during a social story intervention for two children diagnosed with PDD-NOS. *Journal of Developmental and Physical Disabilities, 20*, 217-229. doi:10.1007/s10882-007-9090-4

Dwan, K., Altman, D. G., Arnaiz, J. A., et al. (2008). Systematic review of the empirical evidence of study publication bias and outcome reporting bias. *PLoS One, 3*(8), e3081.

*Dykstra, J. R., Boyd, B. A., Watson, L. R., Crais, E. R., & Baranek, G. T.

(2012). The impact of the advancing social-communication and play (ASAP) intervention on preschoolers with autism spectrum disorder. *Autism: The International Journal of Research & Practice, 16*, 27-44.

Easterbrook, P. J., Berlin, J. A., Gopalan, R., Matthews, D. R. (1991). Publication bias in clinical research. *Lancet, 337,* 867–872.

Faith, M. S., Allison, D. B., & Gorman, B. S. (1996). Meta-analysis of single-case research. In R. D. Franklin, D. B. Allison, & B. S. Gorman (Eds.), *Design and analysis of single-case research* (pp. 245-277). Hillsdale, NJ: Erlbaum.

*Feng, H., Lo, Y., Tsai, S., & Cartledge, G. (2008). The effects of theory-of-mind and social skill training on the social competence of a sixth-grade student with autism. *Journal of Positive Behavior Interventions, 10*, 228-242.

*Ferraioli, S. J., & Harris, S. L. (2011). Teaching joint attention to children with autism through a sibling-mediated behavioral intervention. *Behavioral Interventions, 26,* 261-281. doi:10.1002/bin.336

*Finnigan, E., & Starr, E. (2010). Increasing social responsiveness in a child with autism: A comparison of music and non-music interventions. *Autism: The International Journal of Research and Practice, 14*, 321-348.

*Ganz, J. B., Bourgeois, B. C., Flores, M. M., & Campos, B. A. (2008). Implementing visually cued imitation training with children with autism spectrum disorders and developmental delays. *Journal of Positive Behavior Interventions, 10,* 56-66. doi:10.1177/1098300707311388

*Garfinkle, A. N., & Schwartz, I. S. (2002).

interactions in children with autism and other developmental disabilities in inclusive preschool classrooms. *Topics in Early Childhood Special Education, 22*, 26- 38.

*Gena, A. (2006). The effects of prompting and social reinforcement on establishing social interactions with peers during the inclusion of four children with autism in preschool. *International Journal of Psychology, 41*, 541-554.

*Gena, A., Couloura, E., & Kymissis, E. (2005). Modifying the affective behavior of preschoolers with autism using in-vivo or video modeling and reinforcement contingencies. *Journal of Autism and Developmental Disorders, 35*, 545-556.

*Gena, A., Krantz, P. J., McClannahan, L. E., & Poulson, C. L. (1996). Training and generalization of affective behavior displayed by youth with autism. *Journal of Applied Behavior Analysis, 29*, 291-304.

Glass, G. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher, 5*(9), 3-8.

*Gonzalez-Lopez, A., & Kamps, E. M. (1997). Social skills training to increase social interactions between children with autism and their typical peers. *Focus on Autism & Other Developmental Disabilities, 12*, 2-14.

*Gutman, S. A., Raphael, E. I., Ceder, L. M., Khan, A., Timp, K. M., & Salvant, S. (2010). The effect of a motor-based, social skills intervention for adolescents with high-functioning autism: Two single-subject design cases. *Occupational Therapy International, 17*(4), 188-197. doi:10.1002/oti.300

*Hagopian, L. P., Kuhn, D. E., & Strother, G. E. (2009). Targeting social skills deficits in an adolescent with pervasive developmental disorder. *Journal of Applied Behavior Analysis, 42*, 907-5; 911.

*Hancock, T. B., & Kaiser, A. P. (2002). The effects of trainer-implemented Enhanced Milieu Teaching on the social communication of children with autism, *Topics in Early Childhood Special Education, 22*, 33-54.

*Hanley-Hochdorfer, K., Bray, M. A., Kehle, T. J., & Elinoff, M. J. (2010). Social stories to increase verbal initiation in children with autism and asperger's disorder. *School Psychology Review, 39*, 484-9; 492.

*Harper, C. B., Symon, J. B. G., & Frea, W. D. (2008). Recess is time-in: Using peers to improve social skills of children with autism. *Journal of Autism and Developmental Disorders, 38*, 815-12; 826.

Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children, 71*, 165-179.

*Hughes, C., Golas, M., Cosgriff, J., Brigham, N., Edwards, C., & Cashen, K. (2011). Effects of a social skills intervention among high school students with intellectual disabilities and autism and their general education peers. *Research and Practice for Persons with Severe Disabilities, 36*, 46-61.

*Hwang, B. & Hughs, C. (2000). Increasing early social-communicative skills of preverbal preschool children with autism through social interactive training. *Journal of the Association for Persons with Severe Handicaps, 25*, 18-28.

Individuals With Disabilities Education Improvement Act of 2004 (IDEA), Pub. L. No. 108– 446, 118 Stat. 2647 (2004).

*Ingersoll, B., Dvortcsak, A., Whalen, C., & Sikora, D. (2005). The Effects of a Developmental, Social--Pragmatic Language Intervention on Rate of Expressive Language Production in Young Children With Autistic Spectrum Disorders. *Focus on Autism & Other Developmental Disabilities*. *20*, 213-222.

*Ingersoll, B., & Gergans, S. (2007). The effect of a parent-implemented imitation intervention on spontaneous imitation skills in young children with autism. *Research in Developmental Disabilities, 28*, 163-175.

*Ingersoll, B., Lewis, E., & Kroman, E. (2007). Teaching the imitation and spontaneous use of descriptive gestures in young children with autism using a naturalistic behavioral intervention. *Journal of Autism and Developmental Disorders, 36*, 487-505.

*Ingersoll, B., & Schreibman, L. (2006). Teaching reciprocal imitation skills to young children with autism using a naturalistic behavioral approach: Effects on language, pretend play, and joint attention. *Journal of Autism and Developmental Disorders, 37*, 1446-1456.

*Isaksen, J., & Holth, P. (2009). An operant approach to teaching joint attention skills to children with autism. *Behavioral Interventions, 24*, 215-236.

*Jahr, E., Eldevik, S., & Eikeseth, S. (2000). Teaching children with autism to initiate and sustain cooperative play. *Research in Developmental Disabilities, 21*, 151-169.

Jenson, W. R., Clark, E., Kircher, J. C., & Kristjansson, S. D. (2007) Statistic reform: evidence-based practice, meta-analyses, and single subject designs. *Psychology in Schools, 44*, 483-493.

*Jerome, J., & Sturmey, P. (2008). Reinforcing efficacy of interactions with preferred and nonpreferred staff under progressive-ratio schedules. *Journal of Applied Behavior Analysis, 41*, 221-225. doi:10.1901/jaba.2008.41-221

*Jones, E. A. (2009). Establishing response and stimulus classes for initiating joint attention in children with autism. *Research in Autism Spectrum Disorders, 3*, 375-15; 389.

*Jones, E. A., Carr, E. G., & Feeley, K. M. (2006). Multiple effects of joint attention intervention for children with autism. *Behavior Modification, 30*, 782-834.

*Jones, E. A., Feeley, K. M., & Takacs, J. (2007). Teaching spontaneous responses to young children with autism, *Journal of Applied Behavior Analysis, 40*, 565-570.

*Jull, S., & Mirenda, P. (2011). Parents as play date facilitators for preschoolers with autism. *Journal of Positive Behavior Interventions, 13*, 17-14; 30.

*Jung, S., Sainato, D. M., & Davis, C. A. (2008). Using high-probability request sequences to increase social interactions in young children with autism. *Journal of Early Intervention, 30*, 163-25; 187.

*Jurgens, A., Anderson, A., & Moore, D. W. (2009). The effect of teaching PECS to a child with autism on verbal behaviour, play, and social functioning. *Behaviour Change, 26*, 66-81.

Kanner L. *(1949).* Problems of nosology and psychodynamics of early infantile autism. *American Journal of Orthopsychiatry, 19*, 416-26.

Kasari, C., Freeman, S., Paparella, T., Wong, C., Kwon, S., & Gulsrud, A. (2005). Early intervention in autism: Focus on core deficits. *Clinical Neuropsychiatry, 2*, 380-388.

Kasari, C., & Rotheram-Fuller, E. (2007). Peer relationships of children with autism: Challenges and interventions. In E. Hollander & E. Anagnostou (Eds.) *Clinical manual for the treatment of autism*. Arlington, VA: American Psychiatric Publishing, Inc.

Kavale, K. A., & Moster, M. P. (2004). Social skills interventions for individuals with learning disabilities. *Learning Disability Quarterly, 27*, 31-43.

*Kern, P., & Aldridge, D. (2006). Using embedded music therapy interventions to support outdoor play of young children with autism in an inclusive community-based child care program. *Journal of Music Therapy, 43*, 270-292.

*Kleeberger, V., & Mirenda, P. (2010). Teaching generalized imitation skills to a preschooler with autism using video modeling. *Journal of Positive Behavior Interventions, 12*, 116-127.

*Koegel, R. L., Vernon, T. W., & Koegel, L. K. (2009). Improving social initiations in young children with autism using reinforcers with embedded social interactions. *Journal of Autism and Developmental Disorders, 39*, 1240-1251.

*Koegel, R. L., Werner, G. A., Vismara, L. A., & Koegel, L. K. (2005). The effectiveness of contextually supported play date interactions between children with autism and typically developing peers. *Research and Practice for Persons with Severe Disabilities. 30*, 93-102.

*Kohler, F. W., Anthony, L.J., Stighner, S.A., & Hoyson, M. (2001). Teaching social interaction skills in the integrated preschool: An examination of naturalistic tactics. *Topics in Early Childhood Special Education, 21*, 93-103.

*Kohler, F. W., Greteman, C., Raschke, D., & Highnam, C. (2007). Using a buddy skills package to increase the social interactions between a preschooler with autism and her peers. *Topics in Early Childhood Special Education*, *27*, 155-163.

*Kohler, K. W., Strain. P. S., Hoyson, M., Davis, L., Donina, W. M., & Rapp, N. (1995). Using a group-oriented contingency to increase social interactions between children with autism and their peers- a preliminary-analysis of corollary supportive behaviors. *Behavior Modification, 19*, 10-32.

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

*Kravits, T. R., Kamps, D. M., Kemmerer, K., & Potucek, J. (2002). Brief report: Increasing communication skills for an elementary-aged student with autism using the picture exchange communication system. *Journal of Autism and Developmental Disorders, 32*, 225-230.

*Krebs, M. L., McDaniel, D. M., & Neeley, R. A. (2010). The effects of peer training on the social interactions of children with autism spectrum disorders. *Education, 131*, 393-403.

*Kuhn, L. R., Bodkin, A. E., Devlin, S. D., & Doggett, R. A. (2008). Using pivotal response training with peers in special education to facilitate play in two children with autism. *Education and Training in Developmental Disabilities, 43,* 37-45.

*Lacava, P. G., Rankin, A., Mahlios, E., Cook, K., & Simpson, R. L. (2010). A single case design evaluation of a software and tutor intervention addressing emotion recognition and social interaction in four boys with ASD. *Autism: The International Journal of Research and Practice, 14,* 161-178.

*Laushey, K. M., & Heflin, L. J. (2000). Enhancing social skills of kindergarten children with autism through the training of multiple peers as tutors. *Journal of Autism and Developmental Disorders, 30,* 183-193.

*Laushey, K. M., Heflin, L. J., Shippen, M., Alberto, P. A., & Fredrick, L. (2009). Concept mastery routines to teach social skills to elementary children with high functioning autism. *Journal of Autism and Developmental Disorders, 39,* 1435-1448.

*Leaf, J. B., Dotson, W. H., Oppeneheim, M. L., Sheldon, J. B., & Sherman, J. A. (2010). The effectiveness of a group teaching interaction procedure for teaching social skills to young children with a pervasive developmental disorder. *Research in Autism Spectrum Disorders, 4,* 186-198. doi:10.1016/j.rasd.2009.09.003

*Leaf, J. B., Taubman, M., Bloomfield, S., Palos-Rafuse, L., Leaf, R., McEachin, J., & Misty L Oppenheim. (2009). Increasing social skills and pro-social behavior for three children diagnosed with autism

through the use of a teaching package. *Research in Autism Spectrum Disorders, 3*, 275-289.

*Lee, R., McComas, J. J., & Jawor, J. (2002). The effects of differential and lag reinforcement schedules on varied verbal responding by individuals with autism. *Journal of Applied Behavior Analysis, 35*, 391-402.

*Lee, R., & Sturmey, P. (2006). The effects of lag schedules and preferred materials on variable responding in students with autism. *Journal of Autism and Developmental Disorders, 36*, 421-428.

*Liber, D. B., Frea, W. D., & Symon, J. B. G. (2008). Using time-delay to improve social play skills with peers for children with autism. *Journal of Autism and Developmental Disorders, 38*, 312-323.

*Licciardello, C. C., Harchik, A. E., & Luiselli, J. K. (2008). Social skills intervention for children with autism during interactive play at a public elementary school. *Education and Treatment of Children, 31*, 27-37.

*Loftin, R. L., Odom, S. L., & Lantz, J. F. (2008). Social interaction and repetitive motor behaviors. *Journal of Autism and Developmental Disorders, 38*, 1124-1135.

Lord, C., Wagner, A., Rogers, S., Szatmari, P., Aman, M., Charman, T., et al. (2005). Challenges in evaluating psychosocial interventions for autistic spectrum disorders. *Journal of Autism and Developmental Disorders*. *35*, 695-708.

*MacDonald, R., Sacramone, S., Mansfield, R., Wiltz, K., & Ahearn, W. H. (2009). Using video modeling to teach reciprocal pretend play to children with autism. *Journal of Applied Behavior Analysis, 42*, 43-55.

*Maione, L., & Mirenda, P. (2006). Effects of video modeling and video feedback on peer-directed social language skills of a child with autism. *Journal of Positive Behavior Interventions, 8,* 106-118.

*Martins, M. P., & Harris, S. L. (2006). Teaching children with autism to respond to joint attention initiations. *Child & Family Behavior Therapy, 28,* 51-68.

*Marzullo-Kerth, D., Reeve, S. A., Reeve, K. F., & Townsend, D. B. (2011). Using multiple-exemplar training to teach a generalized repertoire of sharing to children with autism. *Journal of Applied Behavior Analysis, 44,* 279-294. doi:10.1901/jaba.2011.44-279

*Matson, J. L. & Francis, K. L. (1994). Generalizing spontaneous language in developmentally delayed children via a visual cue procedure using caregiver as therapists. *Behavior Modification, 18,* 186-197.

Matson, J. L., Matson, M. L., & Rivet, T. T. (2007). Social-skills treatments for children with autism spectrum disorders an overview. *Behavior Modification, 31,* 682-707.

*McDonald, M. E., & Hemmes, N. S. (2003). Increases in social initiation toward an adolescent with autism: reciprocity effects. *Research in Developmental Disabilities, 24,* 453-465.

Miller, A. L. (2005). *Interventions targeting reciprocal social interaction in children and young adults with autism spectrum disorders: A meta-analysis*. Unpublished dissertation, University of Utah.

*Mitchel, K., Regehr, K., Reaume, J., & Feldman, M. (2010). Group social skills training for adolescents with asperger syndrome or high functioning autism. *Journal on Developmental Disabilities, 16,* 52-63.

*Mundschenk, N. A., & Sasso. G. S. (1995). Assessing sufficient social exemplars for students with autism. *Behavioral Disorders, 21*, 62-78.

National Research Council (NRC). (2001). *Educating children with autism.* Washington, DC: National Academy Press.

*Nikopoulos, C. K., & Keenan, M. (2003). Promoting social initiation in children with autism using video modeling. *Behavioral Interventions, 18*, 87-108.

*Nikopoulos, C. K., & Keenan, M. (2004). Effects of video modeling on social initiations by children with autism. *Journal of Applied Behavior Analysis, 37*, 93-96.

*Norris, C., & Dattilo, J. (1999). Evaluating effects of a social story intervention on a young girl with autism. *Focus on Autism and other Developmental Disabilities, 14*, 180-186.

*Owen-DeSchryver, J. S., Carr, E. G., Cale, S. I., & Blakeley-Smith, A. (2008). Promoting social interactions between students with autism spectrum disorders and their peers in inclusive school settings. *Focus on Autism and Other Developmental Disabilities, 23*, 15-28.

*Ozdemir, S. (2008). Using multimedia social stories to increase appropriate social engagement in young children with autism. *Turkish Online Journal of Educational Technology, 7*, 80-88.

Parker, R., & Hagan-Burke, S. (2007). Useful effect sizes interpretations for single case research. *Behavior Therapy, 38*, 95-105.

Parker, R. I., Vannest, K. J., & Brown, L. (2009). The improvement rate difference for single-case research. *Exceptional Children, 75*, 135-150.

*Pierce, K., & Schreibman, L. (1997). Multiple peer use of pivotal response training to increase social behaviors of classmates with autism: Results from trained and untrained peers. *Journal of Applied Behavior Analysis, 30*, 157-160.

*Randolph, J. K., Stichter, J. P., Schmidt, C. T., & O'Connor, K. V. (2011). Fidelity and effectiveness of PRT implemented by caregivers without college degrees. *Focus on Autism and Other Developmental Disabilities, 26*, 230-238.

Rao, P. A., Beidel, D. C., & Murray, M. J. (2008). Social skills interventions for children with Asperger's syndrome or high-functioning autism: a review and recommendations. *Journal of Autism and Developmental Disorder, 38*, 353-361.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical Linear Models* (2nd ed.). Newbury Park, CA: Sage Publications.

Raudenbush, S. W., Bryk, A. S., Cheong, Y. F., & Congdon Jr., R. T. (2004). *HLM 6: Hierarchical linear and nonlinear modeling*. Lincolnwood, IL: Scientific Software International.

*Reagon, K. A., & Higbee, T. S. (2009). Parent-implemented script fading to promote play-based verbal initiations in children with autism. *Journal of Applied Behavior Analysis, 42*, 659-664. doi:10.1901/jaba.2009.42-659

*Reeve, S. A., Reeve, K. F., Townsend, D. B., & Poulson, C. L. (2007). Establishing a generalized repertoire of helping behavior in children with autism. *Journal of Applied Behavior Analysis, 40*, 123-136.

*Reichow, B., & Sabornie, E. J. (2009). Brief report: Increasing verbal greeting initiations for a student with autism via a social StoryTM

intervention. *Journal of Autism and Developmental Disorders, 39,* 1740-1743.

Reichow, B., & Volkmar, F. R. (2010). Social skills interventions for individuals with autism: Evaluation for evidence-based practices within a best evidence synthesis framework. *Journal of Autism and Developmental Disorders, 40,* 149-166.

Reichow, B., Volkmar, F. R., & Cicchetti, D. V. (2008). Development of an evaluative method for determining the strength of research evidence in autism. *Journal of Autism and Developmental Disorders, 38,* 1311-1319.

Richards, S. B., Taylor, R. L., Ramasamy, R., & Richards, R. Y. (1999). *Single subject research application in educational and clinical settings*. (pp.134-147 & 145-164). London: Singular Publishing Groups, Inc.

Quinn, M. M., Kavale, K. A., Mathur, S. R., Rutherford, R. B., Jr., & Forness, S. R. (1999). A meta-analysis of social skill interventions for students with emotional or behavioral disorders, *Journal of Emotional and Behavioral Disorders, 7,* 54-64.

*Sancho, K., Sidener, T. M., Reeve, S. A., & Sidener, D. W. (2010). Two variations of video modeling interventions for teaching play skills to children with autism. *Education & Treatment of Children, 33*, 421-442.

*Sansosti, F. J., & Powell-Smith, K. A. (2008). Using computer-presented social stories and video models to increase the social communication skills of children with high-functioning autism spectrum disorders. *Journal of Positive Behavior Interventions, 10,* 162-178.

*Sansosti, F. J., & Powell-Smith, K.A. (2006). Using social stories to improve the social behavior of children with Asperger syndrome. *Journal of Positive Behavior Interventions, 8,* 43-57.

*Scattone, D. (2008). Enhancing the conversation skills of a boy with asperger's disorder through social stories (TM) and video modeling. *Journal of Autism and Developmental Disorders, 38,* 395-400. doi:10.1007/s10803-007-0392-2

Scattone, D. (2007). Social skills interventions for children with autism. *Psychology in the Schools, 44,* 717-726.

*Schrandt, J. A., Townsend, D. B., & Poulson, C. L. (2009). Teaching empathy skills to children with autism. *Journal of Applied Behavior Analysis, 42,* 17-16; 32.

*Schertz, H. H., & Odom, S.L. (2007). Promoting joint attention in toddlers with autism: A parent-mediated developmental model. *Journal of Autism and Developmental Disorders, 37,* 1562-1575.

Scotti, J. R., Evans, L. M., Meyer, L. H., & Walker, P. (1991). A meta-analysis of intervention research with problem behavior: Treatment validity and standards of practice. *American Journal of Mental Retardation, 96,* 233-256.

Scruggs, T. E., Mastropieri, M. A., & Casto, G. (1987). The quantitative synthesis of single-subject research: Methodology and validation. *Remedial and Special Education, 8*(2), 24-33.

*Sigafoos, J., Green, V. A., Payne, D., Son, S., O'Reilly, M., & Lancioni, G. E. (2009). A comparison of picture exchange and speech-generating

devices: Acquisition, preference, and effects on social interaction. *AAC: Augmentative and Alternative Communication, 25*, 99-109.

*Simpson, A., Langone, J., & Ayres, K. (2004). Embedded video and computer-based instruction to improve social skills for students with autism. *Education and Training in Developmental Disabilities, 39*, 240-252.

Smith, T., Scahill, L., Dawson, G., Guthrie, D., Lord, C., Odom, S., Rogers, S., & Wagner, A. (2007). Designing research studies on psychosocial interventions in autism. *Journal of Autism and Developmental Disorder*, *37*, 354-366.

*Stephens, C. E. (2008). Spontaneous imitation by children with autism during a repetitive musical play routine. *Autism: The International Journal of Research and Practice, 12*, 645-671.

*Stevenson, C. L., Krantz, P. J., & McClannahan, L. E. (2000). Social interaction skills for children with autism: A script-fading procedure for nonreaders. *Behavioral Interventions, 15*, 1-20.

*Strain, P. S., & Danko, C.D. (1995). Caregivers encouragement of positive interaction between preschoolers with autism and their siblings. *Journal of Emotional and Behavioral Disorders*, *3*, 2-12.

*Strain, P. S., Danko, C. D., & Kohler, F. (1995). Activity engagement and social-interaction development in young-children with autism- an examination of “free” intervention effects, *Journal of Emotional and Behavioral Disorders, 3*, 108-123.

*Strain, P. S. & Kohler, F. W. (1995). Analyzing predictors of daily social skill performance, *Behavioral Disorders. 21*, 79-88.

*Strain, P. S., Kohler, F.W., Storey, K., & Danko, C.D. (1994). Teaching preschoolers with autism to self-monitor their social interaction: an analysis of results in home and school settings, *Journal of Emotional and Behavioral Disorders, 2*, 78-88.

*Taylor, B. A., & Hoch, H. (2008). Teaching children with autism to respond to and initiate bids for joint attention. *Journal of Applied Behavior Analysis, 41*, 377-391. doi:10.1901/jaba.2008.41-377

*Tetreault, A. S., & Lerman, D. C. (2010). Teaching social skills to children with autism using point-of-view video modeling. *Education and Treatment of Children, 33*, 395-419.

*Thiemann, K. S., & Goldstein, H. (2001). Social story, written text cues, and video feedback: effects on social communication of children autism. *Journal of Applied Behavior Analysis, 34*, 425-446.

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

Thornton, A, & Lee, P. (2000). Publication bias in meta-analysis: its causes and consequences. *Journal of Clinical Epidemiology, 53*, 207-216.

*Thorp, D. M., Stahmer, A.C., & Schreibman, L. (1995). Effects of socialdramatic play training on children with autism, *Journal of Autism and Developmental Disorders, 25*, 265-282.

*Tsao, L. L., & Odom, S.L. (2006). Sibling-mediated social interaction intervention for young children with autism. *Topics in Early Childhood Special Education, 26*, 106-123.

U.S. Department of Education. (2002). No Child Left Behind Act of 2001. Retrieved March 29, 2009, from http://www.ed.gov/policy/elsec/leg/esea02/107-110.pdf.

Van den Noortgate, W., & Onghena, P. (2003a). Combining single-case experimental data using hierarchical linear models. *School Psychology Quarterly, 18*, 325-346.

Van den Noortgate, W., & Onghena, P. (2003b). Hierarchical linear models for the quantitative integration of effect sizes in single-case research. *Behavior Research Methods, Instrument, & Computer, 35*, 1-10.

Van den Noortgate, W., & Onghena, P. (2007). The aggregation of single-case results using hierarchical linear models. *The Behavior Analyst Today, 8*, 196-208.

*Vismara, L. A., & Lyons, G. L. (2007). Using perseverative interests to elicit joint attention behaviors in young children with autism: Theoretical and clinical implications for understanding motivation. *Journal of Positive Behavior Interventions, 9*, 214-228.

*Vismara, L. A., & Rogers, S. J. (2008). The Early Start Denver Model: A case study of an innovative practice. *Journal of Early Intervention, 31*, 91-108.

Wang, S.Y., Cui, Y., & Parrila, R. (2011). Examining the effectiveness of peer-mediated and video-modeling social skill interventions for children with autism spectrum disorders: A meta-analysis of single-case research Using HLM. *Research in Autism Spectrum Disorders, 5*, 562-569.

Wang, S.Y., & Parrila, R. (2008). Quality Indicators for single–case research on social skill interventions for children with autistic spectrum disorder. *Developmental Disabilities Bulletin, 36,* 81-105.

*Whalen, C., & Schreibman, L. (2003). Joint attention training for children with autism using behavior modification procedures. *Journal of Child Psychology and Psychiatry. 44*, 456-468.

World Health Organization. (1992). *International classification of diseases: Diagnostic criteria for research* (10th ed.). Geneva, Switzerland: Author.

*Yang, N. K., Schaller, J. L., Huang, T., Wang, M. H., &Tsai, S. (2003). Enhancing appropriate social behaviors for children with autism in general education classrooms: An analysis of six cases. *Education and Training in Developmental Disabilities*, *38*, 405-416.

CHAPTER V

GENERAL DISCUSSION

This dissertation focused on the methodological quality and effectiveness of single-case studies on social skills interventions (SSIs) for individuals with autism spectrum disorder (ASD). The demands for SSIs for individuals with ASD have escalated not only because of the core deficits in social interaction in ASD but also because of the increasing number of children being diagnosed with ASD and the needs for their successful inclusion in the schools. The expansion of diagnostic criteria from autism to ASD and the increasing awareness of ASD have both helped to identify more children with ASD. The prevalence rate has increased from less than 10 to over 100 in every 10000 individuals currently (Matson & Kozlowski, 2011). Individuals with ASD can face greater challenges as the current practice of inclusive education offers more opportunities to students with disabilities to interact with peers. With the expanding demand and funding, multifarious models have been proposed and studied. On the other hand, tools and research for the investigation of methodological quality and treatment effectiveness of the intervention studies have not developed at the same rate. This dissertation attempted to fill the gap with three studies focusing specifically on methodological and meta-analysis of the SSI studies with individuals with ASD.

The first study compiled a number of quality indicators into a checklist that can be used as a tool to identify high-quality SSIs that can inform evidence-based practices. The study illustrates a rationale for these quality indicators and demonstrates how the quality checklist can be applied to single-case research on SSIs for individuals with ASD. The second study

adopted a meta-analysis method, Hierarchical Linear Modeling (HLM), to investigate and compare treatment effects of two common intervention approaches, peer-mediated and video-modeling interventions, and to examine the impact of moderating factors on intervention outcomes of these studies. Both approaches were found to improve the social behavior of individuals with ASD significantly. The third study applied HLM to the investigation of treatment effectiveness of 115 SSI studies using varied intervention methods and investigated the impact of different moderators across the levels of dependent variables, participants, and studies. In Study Two, age emerged as a significant moderator of the treatment effect; more improvement was noticed in SSIs with younger children with ASD. However, age was not a significant moderator in Study Three. Additionally, Study Three found that the adoption of multiple-baseline or reversal design was associated with bigger treatment effects. Using the quality checklist and meta-analysis method developed in these studies, people who are interested in this field can identify the empirically supported SSIs and examine critical factors for evidence-based practice.

The examination of methodological quality and treatment effectiveness are central to the research synthesis of intervention studies. However, a few issues related to the examination of these two aspects warrant further discussion. First of all, there are concerns regarding whether methodological quality and intervention effects should have their order of priority. The studies with poor methodological quality have more threats to internal validity, and these studies are more likely to be excluded from further investigation and less attention is paid to their effects. However, publication bias arises if we exclude

the papers with poor methodological quality and only examine the treatment outcomes of the studies with good quality (Reichow, 2011; Reichow & Volkmar, 2010). Hence, a comprehensive examination that simultaneously investigates both methodological quality and intervention effectiveness of all targeted studies may be helpful to minimize the bias and represent the full picture of SSIs. Additionally, a comprehensive review that covers the studies with poorer methodological quality can inform us of the common mistakes of research implementation and remind researchers to make proper corrections. As well, there will be fewer chances to miss promising intervention methods in a comprehensive report because the studies with poorer rating on quality indicators but showing positive outcomes are more likely to be included. The report can encourage the researchers who are interested in those particular interventions to improve research methodology and examine the treatment effectiveness again. Additionally, it may impact editors decisions about papers that meet quality guidelines but report minimal effectiveness, and thus minimize the publication bias of favouring studies with significant positive outcomes (Dwan, Altman, Arnaiz, et al., 2008; Easterbrook, Berlin, Gopalan, & Matthews, 1991). The intervention studies with proven research quality but minimal effectiveness should not be excluded from the examination as these studies can inform us of the ineffective intervention methods and warn against further waste of time and resources.

Secondly, quality indicators and meta-analysis methods used for the review of SSI studies for individuals with ASD should be readjusted based on the diverse focus of different phases of research development. Smith et al. (2007) suggested that there should be different phases with different focuses

for research development of SSI studies for individuals with ASD, ranging from initial efficacy studies of new intervention techniques to final effectiveness studies of community application of the promising methods. To reflect differences between the phases, the tools used to review SSI research with individuals with ASD should be accommodated accordingly. For example, the development and implementation of review tools to examine methodological quality and treatment effects of single-case SSI studies should be targeted at the initial phase because single-case research is more likely to be adopted by new intervention techniques at this phase. At this phase, the SSI review should include the development of examination tools such as a list of quality indicators that can help identify single-case studies with proven quality as well as be used as guidelines for improving methodological quality of future studies. The second focus of this phase may include the development of the synthesis tools of treatment effectiveness suitable for examination of single-case SSIs, so that we can properly investigate the intervention outcomes across various new methods and programs.

At the next phase, research development of SSIs may focus on the maintenance and generalizability of the SSIs that have demonstrated initial efficacy through a series of single-case studies. Therefore, synthesis tools need to be modified in order to target on the examination of the maintenance and generalizability of these SSIs using single-case research. Maintenance and generalizability are vital to intervention success for individuals with ASD (Bellini & Peters, 2008; National Research Council [NRC], 2001). However, the studies that examine maintenance or generalizability of the effects may fail to pass some quality criteria, such as consistency of settings and materials,

which were important and included in the research review at the initial phase. As a result, modification of quality indicators is required when we investigate the intervention studies that aim to test the maintenance and generalizability of the SSIs using single-case studies.

Further modification of synthesis tools will be required when research development of SSIs for individuals with ASD move from the phase that accumulates evidences through single-case studies to examining effectiveness with quasi-experimental studies. The synthesis tools, including quality indicators and meta-analysis methods, used for single-case research and quasi-experimental research are substantially different. The synthesis tools should be modified to be appropriate and applicable to the examination of methodological quality and treatment effects of quasi-experimental research. Subsequently, synthesis tools may require further modification if they are used to investigate the evidence for maintenance and generalizability effects of SSIs from quasi-experimental studies.

Special attention should be paid to the use of effect sizes in the interpretation of treatment effectiveness. Parker and Hagan-Burke (2007) and Parker, Vannest, and Brown (2009) stressed that effect sizes alone do not represent the treatment effectiveness and applying effect sizes to the summary of treatment effectiveness is context-dependent. Other factors such as study designs, participant characteristics, statistical methods, treatment settings, and target behaviors can lead to variation in effect sizes and must be taken into account (Campbell, 2004; Jenson et al., 2007). For instance, the participants in a set of single-case studies are generally heterogeneous and their intervention goals can be varied, and subsequently a large effect size from a set of

single-case studies does not mean that the same effect can be generalized to other participants with different characteristics or intervention goals. Therefore, the researchers should cautiously attend to the impacts of different factors such as study designs, participant characteristics, statistical methods, treatment settings, and target behavior on meta-analysis outcomes.

As well, benchmarks of effect sizes should be utilized with caution. The benchmarks that are used to interpret effect sizes in a specific discipline, meta-analysis method, or context may not be suitable to interpret effect sizes in other disciplines and contexts, or from different meta-analysis methods (Beeson & Robey, 2006; Kromrey & Foster-Johnson, 1996; Maughan, Christiansen, Jenson, Olympia, & Clark, 2005; Parker & Brossart, 2003). For example, the effect sizes of the reviewed studies in Studies Two and Three ranged from -0.17 to 2.31 and more than 80 percents of them would be classified as large using Cohen's (1988) guideline. However, the benchmarks of 0.2, 0.5, and 0.8 set by Cohen (1988) to identify small, medium, and large effect sizes were based on the analysis of between-group studies in the discipline of social sciences. The effect sizes of single-case studies are more likely to be larger than 0.8. More than 89% of the studies in Study Three had effect sizes larger than 0.8. Similarly, "large" effect sizes ranging from two to 23 were reported in a review of single-case aphasia studies (Robey, Schultz, Crawford, & Sinner, 1999). Beeson and Robey (2006) stated the effect sizes of their reviews of single-case aphasia studies varied depending on the differences of intervention targets. Moreover, Brossart, Parker, Olson, and Mahadevan (2006) argued for different benchmarks for different meta-analysis methods after they noticed the very different effect sizes coming from five

different analysis methods when they applied these methods to the same set of single-case research studies. Therefore, developing benchmarks that can help us to appropriately interpret the effect sizes from SSI single-case research with individuals with ASD obtained by different methods should be given high priority.

Additionally, utilization of effect sizes should not lead to the exclusion or underestimation of the importance of other approaches to evaluating intervention effects. Allison and Gorman (1993), Morgan and Morgan (2009), Parker and Hagan-Burke (2007), and Van den Noortgate and Onghena (2007) have all recommended that visual analysis should be included as a complimentary approach for meta-analysis of single-case research. Parker et al. (2009) supported the combination of visual analysis and effect sizes because the combination integrates the advantages from both approaches. Effect sizes provide objective and standardized index for the comparison of multiple studies that visual analysis doesn't provide. Yet, visual analysis is more convenient and conservative compared to effect sizes. Similarly, the importance of clinical significance and clinical judgments should not be underestimated (Brossart et al., 2006; Kromrey & Foster-Johnson, 1996). Clinical significance and clinical judgments can reflect the impact of the intervention on a participant's life and can never be replaced by the quantitative values of statistical analysis methods. Therefore, meta-analysis of SSI studies with individuals with ASD should also take visual analysis and clinical judgment into account in order to get a comprehensive picture of treatment effectiveness.

In the meanwhile, additional attention should be paid to the comparison of different synthesis methods after applying these methods to SSI studies with individuals with ASD. There can be different results as meta-analysis methods may interact with study factors including study characteristics, participant characteristics, or data characteristics (i.e., outliers, trends, and the number of data points). For instance, Bellini et al. (2007) adopted the percentage of nonoverlapping data points (PND) to review SSI single-case studies of children with ASD and reported that the outcomes of SSIs were less impressive, but positive outcomes of SSIs were reported in the review done by Reichow and Volkmar (2010) and in Study Three of this dissertation using different meta-analysis methods. Similarly, peer-mediated intervention was found to be effective in Study Two using HLM, but it was not found to be effective in a meta-analysis done by Wang and Spillance (2009) using PND. Different meta-analysis methods may have different limitations when they are applied to SSI single-case studies with individuals with ASD. For example, PND can be less sensitive to detecting the progress in social behavior because there can be high variability in outcome measures of social behavior (Jenson et al., 2007). Therefore, comparing and contrasting various meta-analysis methods as well as visual analysis and clinical significance will help us get a more complete picture of the weaknesses and strengths of each method and choose the synthesis methods that best fit the characteristics of SSI studies of individuals with ASD.

Exploring moderators of intervention effectiveness is critically important to the analysis of intervention studies (Kasari & Rotheram-Fuller, 2007; NRC, 2001), and more systematic investigations of moderators are needed. Different

moderators such as children's characteristics (e.g., IQ, language abilities, severity of disability, and age), intervention strategy and dosage (e.g., duration, frequency, and total hours) have all been argued to impact treatment effects (Bellini & Peters, 2008; Kasari & Rotheram-Fuller, 2007; Magiati, Charman, & Howlin, 2007; NRC, 2001). For instance, Study Two suggests that younger children may improve more in their SSI outcomes. However, age was not found to impact treatment outcomes significantly in Study Three. The issue of insufficient information provided by studies generally hinders the analysis of many possible moderators. As a result, researchers in this field should be encouraged to refer to quality indicators as they implement the studies to improve their research as well as to supply sufficiently detailed information for replication and further analysis. With sufficient information, we can analyze the impacts of varied factors on treatment effectiveness and get a more accurate picture of their importance.

Furthermore, there are other issues for the analysis of moderators' impacts on treatment effectiveness. First of all, the interaction between different moderators may require more attention. For instance, age is likely to interact with intervention approach. Study Two reported that younger children in video-modeling programs improved more in their social behavior but the benefit for younger children at peer-mediated programs was limited. Additionally, the impact of age on treatment effects may not be linear or may interact with other moderators. For instance, younger children with ASD may generally improve more than older children, but older children with ASD at peer-mediated program may gain more because their peers are generally older and can follow the therapist's instructions better and provide better supports

and modeling. Similarly, investigating the relationship between effect sizes and research quality using the combination of quality indicators can be misguided as the impact of each quality indicator can be either positive or negative as shown in Study Three. In order to clarify how each moderator impacts treatment effectiveness or how moderators interact with each other, meta-analyses that focus on homogenous studies (i.e., moderators for video-modeling intervention studies) as well as comparison of moderator effects across these meta-analyses would be helpful. However, clear definitions (e.g., SSI approaches, SSI strategies, target behavior) and implementation and accumulation of homogenous studies of sound quality are prerequisites for progress in the field. As a result, efforts to promote research quality such as the adoption of specific research guidelines must continue.

Finally, ineffective interventions and the related moderators that lead to poor outcomes should not be ignored when we synthesize SSI studies for individuals with ASD. Specifically, proper attention should be given to studies that establish good research quality but demonstrate minimal treatment effects. These studies are important as they may help us identify some popular but unproductive intervention approaches as well as factors that may lead to poor outcomes. With better understanding of ineffective interventions and the factors associated with poor outcomes, there will be less waste of energy and resources on interventions that have been proven ineffective and the implementation of empirically supported interventions will be more likely.

Limitations

Some limitations with the three studies reported in this dissertation should be noted and addressed in future studies. Quality indicators used in

Study Three were the primary indicators that could be scored consistently. The limited variability in these indicators undermined their ability to moderate treatment effects. Additional modification of the quality checklist may be required in order to examine other indicators that can be important to the research quality of SSIs for participants with ASD, or that can introduce systematic bias to the results. For instance, indicators such as cost- and time-effectiveness play an important role in intervention applicability and should be added to the checklist when operational definitions are available and papers report relevant information. Furthermore, the scoring of quality checklist does not reflect the fact that not all indicators are equally important for internal validity. Supplementary endeavors should be made to differentially weight the quality indicators for the scoring of quality checklist to represent the overall study quality more faithfully. For instance, the quality indicators that are crucial to internal validity of single-case research, such as the management of confounding factors and the fidelity of implementation, may deserve more weight.

As more quality indicators are examined, some based on inconsistently reported information, the inter-rater reliability also needs to be examined. Inter-rater reliability was not examined in Study Three as the quality indicator checklist used in Study Three was designed to be as unambiguous as possible and focus on information that could be scored consistently. Establishing inter-rater reliability is recommended in future studies with expanded checklists.

In Study Two and Three, social skills were treated as homogenous. However, the target behaviors of SSIs for individuals with ASD vary widely,

and treatments may also vary in their effectiveness across different behaviors. For example, teaching individuals with ASD to respond to a request can be easier achieved than teaching them to play collaboratively. Future studies should examine treatment effectiveness across different categories of social behavior once there are clearer definitions for these varied social skills and behaviors.

Studies Two and Three are subject to publication bias and limited generalizability of the meta-analysis outcomes because of the study selection and exclusion criteria, restricted number of included intervention studies, and the concerns for order effects of intervention phases. Furthermore, the investigations of moderators' impact were incomplete in Studies Two and Three because we could not collect sufficient information on many potentially important moderators, such as intervention intensity and participant's abilities, consistently across all studies. As a result, setting guidelines for implementing single-case studies of SSIs for children with ASD is strongly encouraged in an attempt to get better information for further analysis.

Similarly, the analysis of intervention strategy as a moderator was not possible in Study Three due to the overlap of intervention strategies used in many of the reviewed studies and the unavailability of a clear distinction between several intervention methods. However, the analysis of intervention methods and strategies as moderators is a very important task for future studies once these barriers can be overcome. Once we have a reliable way to classify intervention methods, we can analyze their impact separately as well as the possible interaction of intervention methods with other moderators such as age or targeted social behavior.

HLM was chosen for the meta-analysis in Studies Two and Three because HLM is a more powerful tool to manage specific data problems related to single-case research including different scales across multiple studies, autocorrelation, trend, and short number of data points. Nevertheless, HLM has limitations with the analysis of nominal variables that contain three or more categories (e.g., multiple intervention approaches) due to the difficulties with coding assignment. To facilitate the analysis of multiple dimensions of one or more moderators, technical issues for coding these variables using HLM should be explored in the future. Finally, the last two studies suggested that applying existing benchmarks for effect sizes developed for experimental studies may not be appropriate. Investigating and developing different standards in future research is required for meaningful interpretation of effect sizes from meta-analyses of SSI studies with individuals with ASD. Different meta-analysis methods, research designs, participants' characteristics, and target behaviors may also require different benchmarks, as discussed above.

Conclusions

This dissertation included three studies that explored the methodological quality and treatment effectiveness of single-case research on SSIs for individuals with ASD. The results of these studies demonstrate the applicability of the methodological quality checklist and HLM. As well, video-modeling and peer-mediated approaches were found to be effective in improving the social behavior of participants with ASD. Overall, positive outcomes were noticed while we applied HLM to the examination of treatment effectiveness of 115 single-case SSI studies with individuals with ASD. Age

was not consistently found to impact treatment effectiveness across Study Two and Study Three. Research design was identified as a moderator in Study Three. Although more efforts should be placed in this field in order to get a comprehensive picture of research quality and effectiveness of single-case SSI studies with individuals with ASD, these three studies can facilitate evidence-based practice through offering an in-depth examination of single-case studies of SSIs for children with ASD. The tools and results from these three studies can help parents, practitioners, researchers and policy-makers to know more about the methodological quality of the studies, overall effectiveness of the SSIs, and what may be the important moderators. As a result, people who are interested in this field have more information to choose, implement, develop, and fund effective SSIs for individuals with ASD in order to facilitate their development.

References

Allison, D. B., & Gorman, B. S. (1993). Calculating effect sizes for meta-analyses: The case of the single case. *Behavior Research and Therapy, 31*, 621-631.

Beeson, P. M., & Robey, R. R. (2006). Evaluating single-subject treatment research: Lessons learned from the aphasia literature. *Neuropsychology Review 16*, 161-169.

Bellini, S., Peters, J. K., Benner, L., & Hopf, A. (2007). A meta-analysis of school-based social skills interventions for children with autism spectrum disorders. *Remedial and Special Education, 28*, 153-162.

Brossart, D. F., Parker, R. I., Olson, E. A., & Mahadevan, L. (2006). The relationship between visual analysis and five statistical analyses in a simple AB single-case research design. *Behavior Modification, 30*, 531-563.

Campbell, J. M. (2004). Statistical comparison of four effect sizes for single-subject designs. *Behavior Modification, 28*, 234-246.

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

Dwan, K., Altman, D. G., Arnaiz, J. A., et al. (2008). Systematic review of the empirical evidence of study publication bias and outcome reporting bias. *PLoS One, 3*(8), e3081.

Easterbrook, P. J., Berlin, J. A., Gopalan, R., & Matthews, D. R. (1991). Publication bias in clinical research. *Lancet, 337,* 867–872.

Jenson, W. R., Clark, E., Kircher, J. C., & Kristjansson, S. D. (2007) Statistic reform: evidence-based practice, meta-analyses, and single subject designs. *Psychology in Schools, 44*, 483-493.

Kasari, C., & Rotheram-Fuller, E. (2007). Peer relationships of children with autism: Challenges and interventions. In E. Hollander & E. Anagnostou (Eds.) *Clinical manual for the treatment of autism*. Arlington, VA: American Psychiatric Publishing, Inc..

Kromrey, J. D., & Foster-Johnson, L. (1996). Determining the efficacy of intervention: The use of effect sizes for data analysis in single-subject research. *Journal of Experimental Education, 65*, 73-93.

Magiati, I., Charman, T., & Howlin, P. (2007). A two-year prospective follow-up study of community-based early intensive behavioural intervention and specialist nursery provision for children with autism spectrum disorders. *Journal of Child Psychology and Psychiatry, 48*, 803-812.

Matson, J. L., & Kozlowski, A. M. (2011). The increasing prevalence of autism spectrum disorders. *Research in Autism Spectrum Disorders, 5*, 418–425.

Maughan, D. R., Christiansen, E., Jenson, W. R., Olympia, D., & Clark, E. (2005). Behavioral parent training as a treatment for externalizing behaviors and disruptive behavior disorders: A meta-analysis. *School Psychology Review, 34,* 267–286.

Morgan, D. L., & Morgan, R. (2009). Data analysis in single-case research. In *Single-case research methods for the behavioral and health sciences* (pp. 205-231). Thousand Oaks, CA: Sage Publications.

National Research Council (NRC). (2001). *Educating children with autism.* Washington, DC: National Academy Press.

Parker, R. I., & Brossart, D. F. (2003). Evaluating single-case research data: A comparison of seven statistical methods. *Behavior Therapy, 34*, 189-211.

Parker, R., & Hagan-Burke, S. (2007). Useful effect sizes interpretations for single case research. *Behavior Therapy, 38*, 95-105.

Parker, R. I., Vannest, K. J., & Brown, L. (2009). The improvement rate difference for single-case research. *Exceptional Children, 75*, 135-150.

Reichow, B. (2011). Development, Procedures, and Application of the Evaluative Method for Determining Evidence-Based Practices in Autism. In B. Reichow, P. Doehring, D.V. Cichetti, & F. R.Volkmar (Eds), *Evidence-Based Practices and Treatments for Children with Autism*. Boston, MA: Springer.

Reichow, B., & Volkmar, F. R. (2010). Social skills interventions for individuals with autism: Evaluation for evidence-based practices within a best evidence synthesis framework. *Journal of Autism and Developmental Disorders, 40*, 149-166.

Reichow, B., Volkmar, F. R., & Cicchetti, D. V. (2008). Development of an evaluative method for determining the strength of research evidence in autism. *Journal of Autism and Developmental Disorders, 38*, 1311-1319.

Robey, R. R., Schultz, M. C., Crawford, A. B., & Sinner, C. A. (1999). Single-subject clinical-outcome research: designs, data, effect sizes, and analyses. *Aphasiology, 13,* 445–473.

Smith, T., Scahill, L., Dawson, G., Guthrie, D., Lord, C., Odom, S., Rogers, S., & Wagner, A. (2007). Designing research studies on psychosocial interventions in autism. *Journal of Autism and Developmental Disorders, 37,* 354-366.

Van den Noortgate, W., & Onghena, P. (2007). The aggregation of single-case results using hierarchical linear models. *The Behavior Analyst Today, 8,* 196-208.

Wang, S.Y., Cui, Y., & Parrila, R. (2011). Examining the effectiveness of peer-mediated and video-modeling social skill interventions for children with autism spectrum disorders: A meta-analysis of single-case research Using HLM. *Research in Autism Spectrum Disorders, 5,* 562-569.

Wang, S.Y., & Parrila, R. (2008). Quality Indicators for single–case research on social skill interventions for children with autistic spectrum disorder. *Developmental Disabilities Bulletin, 36,* 81-105.

Wang, S. Y., Parrila, R. & Cui, Y. (2012). Meta-Analysis of Social Skills Interventions of Single-Case Research for Individuals with Autism Spectrum Disorders: Results from Three-Level HLM. *Journal of Autism and Developmental Disorders.* Advance online publication. doi: 10.1007/s10803-012-1726-2

Wang, P. & Spillance, A. (2009). Evidence-based social skills Interventions for children with autism: A meta-analysis. *Education and Training in Developmental Disabilities. 44,* 318-342.

Appendix A

The Quality Indicator Checklist for Single-Case Research in ASD

Quality indicators	**%**			
Primary Quality indicators	Yes	No	Part	NA
Participants:				
Gender and age of ASD participant(s) is provided				
Ethnicity information of ASD participant(s) is provided				
Recruiting procedure of ASD participant(s) is explained				
Selection criteria of ASD participant(s) are explained				
Information on relevant abilities (IQ, academic performance, or adaptive skills) provided				
ASD diagnosis made by professionals specialized in ASD				
The study used a standardized instrument for diagnosis				
Detailed information on training & qualifications of interventionists provided				
Detailed information on the recruiting procedure of peers provided				
Detailed information of selection criteria of peers provided				
Settings/materials used for social skill training:				
Information on the settings and materials sufficient for replication				
Potential confounding factors caused by the settings/materials controlled				
Independent Variables:				
IVs described in sufficient detail for replication				
Standardized procedure used for implementation (i.e., manual)				
Researchers controlled the contamination between subjects				
IV implemented at least three times at three different time points				
The researchers assessed the fidelity of implementation				

Appendix A (continued)

Dependent Variables:

DVs were operationally defined

DVs clearly linked to target behaviors

Measuring procedure generated a quantifiable index

DVs measured at least 3 times on each baseline phase

The data on each baseline phase present a stable pattern/trend

DVs measured at least 3 times on each intervention phase

The data on each intervention phase presents a stable pattern/trend

The inter-rater agreement over 80% or Kappa over .60 between raters

The inter-rater agreement collected on at least 20% of sessions

The raters were blind to research

The raters were different from the interventionist

Research Designs: using multiple baseline or reversal design

Secondary Quality indicators

External validity:

The researcher reported data on maintenance effect

The data on generalization of effects collected across different contexts

Social validity:

Data on direct gains (other than DVs) reported

Data on secondary gains caused by intervention reported

Data on consumer satisfaction reported

Qualitative data reported for social importance of change in DVs

IV implementation needs minimal adjustment to natural settings

The research examined SV over extended (3 month later) period

The agents used to establish SV blind to research

The agents used to establish SV adopted from typical contexts

Note: Yes = the study meets the criterion; No = the study does not meet the criterion; Part = the study meets the criterion of this quality indicator partially; NA = the quality indicator is not applicable to the study.