Social Story Effectiveness on Social Interaction for Students with Autism: A Review of the Literature

Muhammed A. Karal and Pamela S. Wolfe
The Pennsylvania State University

Abstract: Social stories frequently have been used to improve the social interaction of students with autism spectrum disorder (ASD). This literature review examines the effectiveness of social story interventions on the social interactions of students with ASD including with whom, where, and what formats have been implemented, as well as the methodological rigor of the research. Findings indicate preliminary evidence to suggest that social stories are a promising intervention to increase social interactions. Future research with methodological rigor is needed to reveal the most effective strategy for developing and implementing social story interventions.

Autism spectrum disorder (ASD) is characterized by deficits in social communication and/or social interaction across multiple contexts (American Psychiatric Association, 2013). Impairments in social interactions serve as a defining characteristic of individuals with ASD; deficits can range from mild to severe and may manifest as issues with integrated verbal and nonverbal communication, poor eye contact and body language, deficits in understanding and use of gestures, lack of facial expressions, and other nonverbal communication issues (American Psychiatric Association, 2013). The limited repertoires of social skills of students with ASD affect the ability to form and maintain friendships (Locke, Ishijima, Kasari, & London, 2010). For example, students with ASD are reported to have fewer friendships but have the desire to engage in relationships with others (Locke et. al., 2010). The importance of social skill development is critical. Such deficits may impede childrens' development and increase the risk of social withdrawal and isolation (Matson, Dempsey & LoVullo, 2009). Children who are socially withdrawn may be rejected by peers thus impeding social inclusion (Kago-

hara et al., 2013). Their isolation likely will affect their overall educational experience (Delano & Snell, 2006).

Lack of appropriate social skills has been identified as a critical component of education programming; the Individuals with Disabilities Education Act (IDEA) has stressed that educational goals for children with ASD should include the development of social skills. Social interaction skills center on increasing the frequency of and competency in social interaction. Skills that may be taught can include initiating or responding to other students (verbally, physically, or gesturally), maintaining conversations, offering to help, asking and answering questions, requesting information from others, and interacting in games (Bellini, Peters, Benner, & Hopf, 2007).

Teachers, professionals, peers, and other adults can be very effective in promoting the social engagement of individuals with ASD by using specific intervention strategies (Barnard-Brak, Ivey-Hatz, Ward, & Wei, 2014). Strategies used have included behavioral interventions, modeling, peer training, pivotal response treatment, and scripting (National Autism Center, 2015). One strategy that is frequently used to modify the social engagement of students with ASD is social stories or story based interventions (National Autism Center, 2015).

Social stories are individualized short stories

Correspondence concerning this article should be addressed to Muhammed A. Karal, Department of Educational Psychology, Counseling, and Special Education, The Pennsylvania State University, University Park, PA 16802. E-mail: mak523@psu.edu

that describe situations, concepts, or social skills designed to increase the quality or quantity of social interactions of individuals having ASD with others. The stories describe and explain the cues in the social contexts as well as provide appropriate responses. The goal of a social story is to share relevant information that includes (but is not limited to) where and when a given situation takes place, who is involved, what is occurring, and why (Gray, 1998). Gray (1995) and other researchers (e.g., Hobson, 1993; Trevarthen, Aitken, Papudi, & Roberts, 1996) have proposed that the effectiveness of social story interventions is most consistent with the "theory of mind" (ToM) (Baron-Cohen, 1995) accounts of autism that suggest that individuals with ASD have difficulty understanding perspectives different from their own (Leslie, 1987). Theory of Mind (ToM) is described by Garfield, Peterson, and Perry (2001) as a broad umbrella term used to denote whatever knowledge guides prepositional attitude attribution and the explanation and prediction of behavior by means of inner states and processes. Difficulty in recognizing the thoughts of others may be problematic for individuals with ASD when interpreting social information.

Social stories are an inherently attractive instructional strategy because they are relatively easy to implement and are reported to be applicable to a wide variety of behaviors (Reynhout & Carter, 2006). Social stories were implemented for increasing social and daily-living skills (Agosta, Graetz, Mastropieri, & Scruggs, 2004; Bledsoe, Myles, & Simpson, 2003) and decreasing behaviors (Crozier & Tincani, 2005; Scattone, Wilczynski, Edwards, & Rabian, 2002).

There is a wide variety of ways that social stories have been constructed. Information in the social story can be presented with text and/or visual components (Gray & Garand, 1993). According to Gray's (2010) social story guidelines, a social story should be individualized and consist of seven types of sentences: *descriptive*, *perspective*, *affirmative*, three types of sentences that *coach* and *partial* sentences. According to Gray (2010), the social story formula is complete when the total number of descriptive, perspective and affirmative sentences divided by the number of sentences that coach is equal or larger than two. In the

story, descriptive, perspective, and affirmative sentences can be partially written. A partial sentence encourages the person with ASD to complete a given statement in terms of his or her perception of what will happen next or to make a guess regarding the responses of another person (Gray, 2010).

There have been concerns ascertaining the efficacy of social stories (Mayton, Menendez, Wheeler, Carter, & Chitiyo, 2013). Efficacy of an intervention can be represented in a number of ways, one of which is calculation of effect size. Effect size serves as an indication of the magnitude of the effect of the intervention. One calculation of effect size commonly used is the percentage of non-overlapping data (PND). PND provides the overlap between treatment and baseline phases by counting the number of data points in the treatment phase that exceed the highest data point of the baseline phases. Several meta-analyses have been conducted by researchers that report effect size. Two meta-analyses were conducted by Reynhout and Carter (2006, 2011) that provided a descriptive synthesis (16 and 62 studies, respectively) using the percentage of non-overlapping data (PND) and three different metrics (PND, percentage of data exceeding the median (PEM), and improvement rate difference (IRD)) to determine the effectiveness of the interventions. Both of these studies questioned the effectiveness of social stories ($M = 43\%$ and $M = 52\%$ respectively). A meta-analysis conducted by Kokina and Kern (2010) that included 18 studies also used PND metric and reported questionable intervention effectiveness ($M = 60\%$). Similarly, in their meta-analysis Test, Richter, Knight, and Spooner (2011) found that the PND scores for social story intervention effectiveness were questionable ($M = 50\%$).

Measures other than PND have been introduced by researchers to examine effect size including percentage of all non-overlapping data (PAND), non-overlap of all pairs (NAP), IRD, and PEM. Of all the established effect size metrics most have problems for applied research such as having esoteric meanings, assuming data properties lacking in datasets, or encouraging oversimplified misinterpretations (Parker, Vannest, & Brown, 2009). Improvement rate difference (IRD) (Parker et

al., 2009) has been suggested to provide a comparison with better sensitivity, allowing for confidence intervals, and providing improvement rates between baseline and treatment phases.

Another component of efficacy is the quality of the research itself. In an effort to assess methodological rigor, the National Autism Center developed a Scientific Merit Rating Scale (SMRS). This scale involves five dimensions including research design, measurement of the dependent variable, measurement of the independent variable, participant ascertainment, and generalization and maintenance effects (NAC, 2015) to evaluate the strength of the intervention efficacy.

The purpose of this review is to examine the literature addressing the efficacy of social stories for improving the social interactions of students with ASD. The review includes analyses related to whom and where social stories have been used, and what format and implementation strategies have been employed. In addition to descriptive analyses, the efficacy of the interventions is reported using IRD metric as well as the SMRS to evaluate methodological rigor.

Method

Search and Selection Procedures

This review focused on studies published from 1993 to 2015. This time period was used because Gray and Garand developed the initial set of guidelines for creating social stories in 1993; accordingly, reviewing the literature of the past three decades permitted an integrated and comprehensive understanding of the literature.

A comprehensive search to locate articles for this literature review was conducted using a three-phase methodology: a key term search, a title and abstract review, and a check of articles' reference lists. First, four electronic databases were searched: Lion Search, Education Resources Information Center (ERIC), PsycINFO, and ProQuest Education Journals. In order to maximize the yield of the database, multiple search terms were used. Intervention-related search terms included: *social story, and social interaction, or social engagement, or socially appropriate behavior,* and population-

related search terms included: *autism, or autism spectrum disorder, or Asperger syndrome, or Asperger's disorder.* Using these search terms, 325 articles were identified initially. The following criteria were used for inclusion in this review: (a) participants of the study were students with ASD, (b) the titles and/or abstracts contained the terms *social interactions, or communication skills, or social engagement, or prosocial behavior, or social communication, or socially appropriate behavior,* (c) the study was published in a peer-reviewed journal, and (d) the study was data-based and experimental. The articles were classified into relevant and non-relevant sets based on these inclusion criteria. Seventy-eight articles were identified; articles about storytelling, narrative therapy, social scripts and autism stories were excluded. In addition, studies based on the teachers' and parents' perceptions and unpublished dissertations were excluded from the review. Third, the reference lists of the chosen articles were reviewed to identify additional studies about social story interventions that also met the criteria. Consequently, 12 peer-reviewed journal articles (Barry & Burlew, 2004; Crozier & Tincani, 2007; Delano & Snell, 2006; Kagohara et al., 2013; Malmberg, Charlop & Gershfeld, 2015; Norris & Dattilo, 1999; Sansotti & Powell-Smith, 2006, 2008; Scattone, 2008; Scattone, Tingstrom, & Wilczynski, 2006; Schneider & Goldstein, 2010; Thiemann & Goldstein, 2001) were identified for the literature review. As an inter observer agreement (IOA) for identified articles regarding if an article met the inclusion criteria, a doctoral student read all articles and confirmed that all studies met the inclusion criteria (100%).

Effect Size

The IRD metric was calculated by determining the exceeding data points in treatment phase over all baseline data points divided by the total number of improved data points in that phase while eliminating overlapping data points between phases. For this review, each data point from baseline and treatment phases were extracted from graphs using the data extraction application WebPlotDigitizer (Rohatgi, 2015). The article pages that contain the graphics were saved as JPEG format files and dragged into the application. After

adjusting the x and y axes of the graph, each data point was spotted on the application. Specified data then were imported into the online calculator developed by Vannest, Parker, and Groen (2011). According to Parker et al. (2009), the maximum IRD score is 1.00, while scores greater than .75 indicate very large effect sizes, scores between .70 and .75 indicate large, scores between .51 and .70 indicate moderate, and scores less than .50 indicate small effect sizes.

Methodological Rigor

The SMRS rating scale was used to evaluate the methodological rigor of studies. Per the SMRS guidelines, each study was evaluated on the five dimensions and assigned a score that ranged between 0 and 5, with 0 representing the poorest experimental rigor and 5 the strongest. The combined formula used to determine methodological rigor was as follows: research design (.30) + dependent variable (.25) + participant ascertainment (.20) + procedural integrity (.15) + generalization and maintenance (.10). Scores that were calculated as 3, 4, and 5 indicated sufficient scientific rigor permitting firm conclusions about the intervention effectiveness while scores of 2 indicated initial evidence with more rigorous research needed; scores of 0 and 1 provide insufficient scientific evidence to state whether any beneficial or harmful effects were evidenced (NAC, 2015). Generalization and maintenance are reported in the current review but an in depth discussion related to these variables are reported in Karal and Wolfe (in progress).

Reliability

Reliability calculations were completed by two raters. Two doctoral students served as independent calculators for IRD metrics and raters for SMRS coding. Reliability checks were completed on 100% of both IRD calculations and SMRS ratings. Each rater independently calculated IRD scores and coded each component for SMRS score reliability. The percentage of inter-rater agreement was calculated by dividing the number of agreement by number of agreement plus disagreement and multiply-

ing by 100. Agreement rates were 91% for IRD metrics and 91.76% for SMRS scores.

Results

A summary of the descriptive information including gender, age, settings, study designs, independent and dependent variables is presented in Table 1. Table 2 provides the formats of the social stories, implementation methods, other strategies and investigation results with mean IRD scores. Table 3 includes the components and total SMRS scores.

Participants and Settings

Twelve peer-reviewed, single-subject studies involved 31 participants (27 males and four females). Thirty participants were identified as having ASD, while one participant was identified as language-impaired. The disability criteria for inclusion in the review were students with ASD. Three out of 22 students who were diagnosed with ASD were identified as having severe autism, and five were identified as having mild to moderate autism. Three students were diagnosed with Asperger syndrome (per DSM-IV categories). Only six of the studies (Kagohara et al., 2013; Malmberg et al., 2015; Sansotti & Powell-Smith, 2008; Scattone, 2008; Schneider & Goldstein, 2010; Thiemann & Goldstein, 2001) provided standardized data about the participants. In most instances, participants were given diagnostic labels of only either autism or ASD.

Four out of 12 studies (Delano & Snell, 2006; Norris & Dattilo, 1999; Scattone et al., 2006; Thiemann & Goldstein, 2001) provided information about participants' reading abilities. The reported reading skills of participants in the studies ranged from those who were able to identify basic sight words (e.g., I, will, James) to those who read independently. Communication skills ranged from individuals who were non-verbal to individuals who were able to verbally communicate. Three participants diagnosed with severe autism were able to repeat any language spoken to them and had receptive language skills. Most of the students who were capable of speech and able to communicate verbally had difficulty with conversation skills and had impaired social communication skills. Eight studies (Crozier

TABLE 1

Summary of the Descriptive Information

Reference	Participants		Setting	Design	Independent Variables	Dependent Variables
	Gender	Age				
Barry & Burlew (2004)	2 (1M and 1F) with severe Autism	7, 8	Play centers in general education classroom	Multiple baseline across participants	(2) Social Stories	Interacting with the materials and/or peers Choice making
Crozier & Tincani (2007)	3 M with Autism	3, 3, 5	Kindergarten classroom	Case study with reversal design	(3) Social Stories	Sitting on the edge of the circular carpet, talking with peers, cooperating and sharing materials
Delano & Snell (2006)	3 M with Autism	6, 6, 9	Play area of a resource classroom, open area between classrooms for	Multiple baseline across participants	(3) Social Stories (Two with pictures)	Saying one or more understandable words, directing gesture or movement toward a peer, responding verbal, gestural or movement response within 5 sec
Kagohara et al., (2013)	2 (1M and 1F) with Asperger Syndrome	10, 10	General education classroom	Multiple baseline across participants	(1) Social Story + (VM)	Partial and full greeting to staff members, researchers, and teachers
Malmberg, Charlop, & Gershfeld (2015)	2 M with Autism	6, 8	University based research clinic	Multiple baseline across participants	(2) Social Stories	Offering to help, reciprocal commenting, empathic congratulatory statements, reciprocal questions
Norris & Dattilo (1999)	1 F with Autism	8	In the work area, outside of her classroom	Case study with reversal design	(3) Social Stories	Initiating or responding to other students verbally, physically, or gesturally
Sansotti & Powell-Smith (2006)	3 M with Autism	10, 11, 9	Fenced in area side of school	Multiple baseline across participants	(3) Social Stories	Sportsmanship behavior Maintaining conversation Joining in

TABLE 1—(*Continued*)

| Reference | Participants | | Setting | Design | Independent Variables | Dependent Variables |
	Gender	Age				
Sansoti & Powell-Smith (2008)	3 M with Autism	6, 8, 9	Fenced in area side of school	Multiple baseline across participants	(3) Social Stories (Computer-Presented + VM)	Maintaining conversation Joining in
Scattone (2008)	1 M with Asperger Syndrome	9	Medical center	Multiple baseline across behaviors	(3) Social Stories + (VM)	Looking at the interaction partner for 3 sec or more, grinning or laughing Any unprompted question or comment to the interaction partner
Scattone, Tingstrom & Wilczynski (2006)	3 M with Autism	8, 8, 13	Classroom, School cafeteria, Area out of classroom	Multiple baseline across participants	(3) Social Stories	Initiating or responding to other students verbally, physically, or gesturally
Schneider & Goldstein (2009)	3 M with Autism	10, 6, 5	Corner of computer room, hallway outside their classroom	Multiple baseline across participants	(3) Social Stories + pictures	Moving away from computer room and walking toward the line at the door, raising his hand and waiting to be called, following directions
Thiemann & Goldstein (2001)	5 M (4 with Autism, 1 LI)	11, 7, 8, 6, 12	Media room in school library	Multiple baseline across behaviors	Social Stories + Written text cues	Securing attention, initiating comments, initiating requests, responding peers

Note. VM = Video Modeling

TABLE 2

Social Story Format, Implementation, Other Strategies and Investigation Results

Reference	Format	Implementation	Other Strategies	Results	Mean IRD
Barry & Burlew (2004)	Written + photos of participants, peers and their classroom NSS	Read by teacher	Verbal/gestural/physical prompting, corrective feedback, verbal praise	Need of prompts for choice making decreased and play minutes increased significantly	1.00
Crozier & Tincani (2007)	Written + pictures SS	Read by teacher	Verbal prompting, Comprehension questions, Mayer-Johnson picture symbols	Prosocial behavior of two students increased Increase reported for other after changing the study design and adding prompting	.54
Delano & Snell (2006)	1 written 2 written + pictures NSS	Read by experimenter	Comprehension questions, peer training	Increase reported for two and gradual increase for one participant in frequency of target skills	.68
Kagohara et al., (2013)	Computer-presented ESS	Viewed and read by student	Verbal prompting	Increases for both participants in frequency of greeting adults	.50
Malmberg, Charlop, & Gershfeld (2015)	Written + pictures NSS	Read by experimenter	Comprehension questions , Verbal prompting, Clipart pictures	Learning criterion was only achieved in the prompting conditions for both participants	.22
Norris & Dattilo (1999)	Written + pictures Randomized use of three different stories NSS	Read by student 4–5 times	–	No increase reported in social interaction, overall decrease in making noises (e.g., television sound effects, gagging, yelling)	.35
Sansotti & Powell-Smith (2006)	Written + pictures SS	Read by student	Verbal prompting, Mayer-Johnson pictures	Increases reported in social engagement for two of the three participants	.62
Sansotti & Powell-Smith (2008)	Computer-presented NSS	Viewed and read by student	Mayer-Johnson pictures, corrective feedback from participants and parents	All students indicated an significant increase in maintaining conversation and joining - in	.71
Scattone (2008)	Written + videotaped Social story SS	Read and viewed by student	Comprehension questions	A significant increase reported for two of the social skills (eye contact, initiations, and smiling)	.80

TABLE 2—(Continued)

Reference	Format	Implementation	Other Strategies	Results	Mean IRD
Scattone, Tingstrom & Wilczynski (2006)	Written SS	Read by teacher to two students Read by student	Comprehension questions	Social skills increase reported for two participants but no change for one participant	.54
Schneider & Goldstein (2009)	Written + pictures ESS	Read by teacher	Comprehension questions, Mayer-Johnson picture symbols	Increase reported for two students in social skills, a slight increase for other	.52
Thiemann & Goldstein (2001)	Written + visual symbols ESS	Read by student	Verbal prompting, Comprehension questions, Self-evaluation using video feedback	Significant increases reported in social skills (getting others' attention initiating comments, requests, and responses)	.87

Note. SS = Social Story provided; ESS = Example of Social Story provided; NSS = No Social Story provided.

& Tincani, 2007; Delano & Snell, 2006; Norris & Dattilo, 1999; Sansotti & Powell-Smith, 2006, 2008; Scattone, 2008; Scattone et al., 2006; Thiemann & Goldstein, 2001) provided information about participants' cognitive abilities, which were typically in the average functioning level.

Challenging behaviors of the participants in the studies included sticking fingers in one's ears to tantrums. Two of the participants from a study (Scattone et al., 2006) exhibited stereotypical behaviors, and three participants in a study conducted by Schneider and Goldstein (2009) had tantrums. Participant descriptions included only diagnostic labels of students.

Three of the 12 studies included 19 peers. Delano and Snell (2006) included six (three boys and three girls) peers who were nominated by their teachers. Three of them were randomly assigned to serve as training peers and play partners during intervention sessions, while the other three acted as novel peers and play partners during play sessions. The research conducted by Sansotti and Powell-Smith (2008) included three similarly-aged peers without disabilities who functioned as models in videos for the participants' social stories and provided peer comparisons. The study by Thiemann and Goldstein (2001) included 10 peers without disabilities who were identified and recommended by each child's regular classroom teacher. Two peers without disabilities participated as social partners with each child to form five triads to help them to read and role play (e.g., practice using props for an upcoming activity).

Ten studies took place in a school setting (Barry & Burlew, 2004; Crozier & Tincani, 2007; Delano & Snell, 2006; Kagohara et al., 2013; Norris & Dattilo, 1999; Sansotti & Powell-Smith, 2006, 2008; Scattone et al., 2006; Schneider & Goldstein, 2010; Thiemann & Goldstein, 2001), one study took place at a medical center (Scattone, 2008), and another study took place in a university based research clinic (Malmberg et al., 2015). There were various types of treatment settings in schools across studies. These setting included general education classrooms ($n = 5$), kindergarten classrooms ($n = 5$), areas outside of their classrooms ($n = 5$), media rooms in school libraries ($n = 5$), and fenced-in areas located on the side of the schools ($n = 6$). Two stu-

TABLE 3

Total SMRS Scores with Components

Reference	Total SMRS Scores	Research Design	DV Measures	Procedural Integrity	Participant Ascertainment	Generalization and Maintenance
Barry & Burlew (2004)	*2*	3	2	2	2	0
Crozier & Tincani (2007)	**4**	3	3	5	4	3
Delano & Snell (2006)	**3**	3	4	4	1	4
Kagohara et al., (2013)	**4**	3	4	5	4	3
Malmberg, Charlop, & Gershfeld (2015)	**4**	3	3	5	5	5
Norris & Dattilo (1999)	**2**	1	2	3	2	0
Sansotti & Powell-Smith (2006)	**4**	3	4	5	5	3
Sansotti & Powell-Smith (2008)	**4**	3	4	5	5	4
Scattone (2008)	**3**	1	4	5	5	3
Scattone, Tingstrom & Wilczynski (2006)	**4**	3	4	5	4	0
Schneider & Goldstein (2009)	**3**	3	4	5	4	0
Thiemann & Goldstein (2001)	**4**	5	3	4	4	4

Note. 3, 4, 5 = Sufficient scientific rigor; 2 = initial evidence with more research needed; 0, 1 = Insufficient scientific rigor.

dents from two different studies (Scattone et al., 2006; Schneider & Goldstein, 2010) participated in different places than did the other participants in their studies due to their specific target behaviors. One treatment setting was in a quiet corner of the school's computer room, and the other was in a school cafeteria after lunch.

Social Story Format and Implementation

Twelve peer-reviewed studies involved 34 social stories for 31 participants in total. There were variations in providing the story and the number of social stories used in the studies. Five out of twelve studies did not provide the stories in the studies (Barry & Burlew, 2004; Delano & Snell, 2006; Malmberg et. al., 2015; Norris & Dattilo, 1999; Sansotti & Powell-Smith, 2008). Social stories were provided in their entirety in the appendices of four studies (Crozier & Tincani, 2007; Sansotti & Powell-Smith, 2006; Scattone, 2008; Scattone et al., 2006). Three studies that were conducted by Kagohara and her colleagues (2013) Sch-

neider and Goldstein (2009) and Thiemann and Goldstein (2001) also provided examples of social stories. The study conducted by Norris and Dattilo (1999) used one social story per participant. Norris and Dattilo (1999) used three different randomly-selected social stories for the participant in order to maintain interest, targeting social interaction only. Scattone (2008) used three stories for three different target social skills for the participant.

In three studies (Kagohara et al., 2013; Scattone, 2008; Sansotti & Powell-Smith, 2008) the social story was presented on a computer. Sansotti and Powell-Smith constructed social stories with a PowerPoint which included similarly-aged peers as video models for the participants. In addition, two stories were presented in written formats, and five stories had pictures or photos to enhance the written information. Only one study (Delano & Snell, 2006) had forms of written stories without pictures and written stories with pictures for different participants in the study. In the study conducted by Barry and Burlew (2004), photographs of participants, peers, and their classrooms were

included alongside the written text. An examination of the sentences revealed that descriptive, sentences that coach, and perspective sentences were present in all of the provided stories, while control, cooperative, and affirmative sentences appeared less frequently.

The social story was read to students or read/viewed by students in all of the studies. In the study conducted by Norris and Dattilo (1999), each participant read the same story two days in a row; the student also read the first and third stories four times and the second story five times. Scattone and her colleagues' study (2006) had two of three participants reading stories twice because the other participant was not able to read fluently; however, he had his story read to him by his teacher. In the studies that had the social stories read to students, four featured teachers as readers (Barry & Burlew, 2004; Crozier & Tincani, 2007; Delano & Snell, 2006; Schneider & Goldstein, 2010).

The social stories were presented immediately prior to the target activity (e.g., talking with peers) for all studies. In the study conducted by Barry and Burlew (2004), the researchers created opportunities for students to participate in play centers after teachers had read social stories to them. The other variation involved comprehension questions to ask about the story. In three studies (Malmberg et. al., 2015; Scattone et al., 2006; Schneider & Goldstein, 2010), teachers asked students questions right after the social story reading. If a student answered incorrectly, the teacher reread the story so the student could respond correctly. In three studies (Barry & Burlew, 2004; Norris & Dattilo, 1999; Thiemann & Goldstein, 2001), the social story remained accessible to the student if he or she wanted to read it again.

Effect Size

The mean effect size was moderate (ES = .61) for social story efficacy on social interaction in 12 of the studies. Table 3 provides the information on mean effect sizes for each study. The authors of four studies (Barry & Burlew, 2004; Sansotti & Powell-Smith, 2008; Scattone, 2008; Thiemann & Goldstein, 2001) reported significant increases in students' social skills and play-per-minutes from baseline to intervention. Mean effect sizes of those studies that reported significant increases are 1.00, .71, .80, and .87, respectively. Four peer-reviewed studies revealed moderate effect sizes, while the other four studies (Malmberg et al., 2015; Norris & Dattilo, 1999; Sansotti & Powell-Smith, 2006; Scattone et al., 2006) showed no change for some participants. The mean effect sizes for those studies were. 22, .35, .62, .54, respectively.

Methodological Rigor

For 12 studies, total SMRS scores ranged from 2 to 4 (M = 3.41). Per components of the scale, the highest quality indicators were related to procedural integrity (M = 4.41), participant ascertainment (M = 3.75), and dependent variable measures (M = 3.41). Lower quality scores occurred for research design (M = 2.83), and generalization and maintenance (M = 2.41). Fewer than five data points in a phase or less than three participants for some studies affected scores related to research design. Four studies that did not include any after treatment data affected scores related to generalization and maintenance.

Two studies (Barry & Burlew, 2004; Norris & Dattilo, 1999) had a total SMRS score of 2; the study conducted by Barry and Burlew (2004) reported positive treatment effects. No study attained the highest SMRS score of 5; however, all nine studies (Crozier & Tincani, 2007; Delano & Snell, 2006; Kagohara et al., 2013; Sansotti & Powell-Smith, 2006, 2008; Scattone, 2008; Scattone et al., 2006; Schneider & Goldstein, 2010; Thiemann & Goldstein, 2001) attained a total SMRS score of 3 or 4; all nine of the studies showed a positive change in the behavior of the participants.

Discussion

Descriptive Analysis

The National Standards Project (NAC) provides essential information about interventions that have been shown to be effective and emphasizes the necessity for evidence-based guidelines for intervention for individuals with ASD (NAC, 2015). It is significantly important for not only educators and individuals with ASD but also parents and service provid-

ers to make a decision about the most beneficial intervention selection. In this review, overall effect sizes indicate that social stories are moderately effective, but specific intervention characteristics are associated with stronger outcomes. According to the NAC, although there are some studies with strong scientific outcomes for an intervention designed for individuals with ASD, additional high quality studies must be conducted to show the effectiveness of the intervention consistently (2015). The NAC standards designate story based interventions, which includes social stories, to be an emerging treatment. As an emerging intervention, social stories related to social interaction need additional and consistent research support to be rated as an established intervention.

Ten out of 12 studies surveyed in this literature review indicated that they drew upon Gray's criteria (developed between 1993 and 2010) to create the social stories. Three studies (Barry & Burlew, 2004; Kagohara et al., 2013; Norris & Dattilo, 1999) did not specifically mention the criteria used but included some references to Gray and Garand (1993). Four studies (Sansotti & Powell-Smith, 2008; Scattone, 2008; Scattone et al., 2006; Thiemann & Goldstein, 2001) showed effective or very effective intervention results, with the exception of one participant in the study conducted by Scattone and her colleagues (2006). All of the other studies used Gray's criteria in developing their social story interventions. On the other hand, five studies (Crozier & Tincani, 2007; Delano & Snell, 2006; Malmberg et al., 2015; Sansotti & Powell-Smith, 2006; Schneider & Goldstein, 2010) that also used Gray's guidelines reported varied results. Moreover, the last two studies' authors reported significant increases (Barry & Burlew, 2004) or slightly positive results (Norris & Dattilo, 1999), but they did not mention using Gray's criteria. Gray's criteria may affect the implementation of social stories, but the adherence to these criteria alone might not result in effective interventions. In the meta-analysis conducted by Reynhout and Carter (2006), social stories that deviated from Gray's suggested ratio (i.e., included more sentences that coach rather than descriptive sentences) seemed to produce better intervention outcomes than did those that followed Gray's criteria (Kokina & Kern, 2010).

There are several points to be made regarding format and implementation as they relate to intervention effectiveness. The four studies (Barry & Burlew, 2004; Sansotti & Powell-Smith, 2008; Scattone, 2008; Thiemann & Goldstein, 2001) that reported significant increases in social interactions all included visual components that differed from the other studies that used visual symbols as printed pictures with the text. These visual components (photographs of participants, peers, and the environment; computer-presented social stories; and video feedback) appeared to be a more effective means of increasing social skills than written text alone or written text with printed pictures. That is to say, these studies' outcomes may have been influenced by their introduction of familiar illustrations or computer usage. Children with ASD appear to remember familiar faces (Volkmar, Sparrow, Rende, & Cohen, 1989). Likewise, computers tend to act as controlled environments with minimal distractions, making the use of computers attractive for the education of children with ASD (Boucenna et al., 2014).

When considering implementation, it should be noted that the number of stories and the readers for the stories varied across studies. The use of one social story for one social situation as opposed to several stories for several situations may affect intervention effectiveness. In relation to readers and reading time, teachers, experimenters, or parents read the stories to students in four studies immediately prior to the target activity. Social stories were slightly more effective when used with students who were able to read, but there was no difference between groups of students with limited and poor reading skills concerning the effectiveness of social stories. These findings imply that, given modifications to their implementation, social stories may be appropriate for students with varying reading skill levels.

There was a wide range of combined strategies in the intervention of social stories. Comprehension questions were assigned in seven studies. Early guidelines identified comprehension as a mandatory component of the intervention to prevent inaccurate interpretation of the situation due to the use of visual

representations (Gray & Garand, 1993), but later guidelines did not mention a mandatory comprehension component. Both assessing comprehension with discussion and questioning are needed because it is important to make certain that individuals with ASD understand the main points of the story (NAC, 2015). The mean effect size for studies that included comprehension questions was moderate ($ES = .59$). Prompting strategies were assigned in four studies for either triggering target behaviors or decreasing challenging behaviors during intervention. The mean effect size of four studies in which prompting strategies were used in conjunction with the intervention was large ($ES = .74$). The other two studies included verbal prompting by creating a verbal prompting only condition. Combining visual cues and verbal cues in a social story may help individuals with ASD understand described skills and behaviors (Dettmer, Simpson, Myles, & Ganz, 2000).

Efficacy

The IRD metric was used in this review to measure the efficacy of the social story interventions. Despite the fact that an effect size only cannot summarize whether the treatment caused the improvement, effect sizes are necessary supplements to visual analysis to establish functional relationship between treatment and outcome (Parker et al., 2009). The IRD metric, in comparison to other established effect size metrics, has already been established in medical research, does not require unwarranted data assumptions, and has obtained confidence intervals (Parker et al., 2009).

Although this review included a small number of interventions, the total average IRD score was slightly higher than that of previous analyses related to social story effectiveness. Social story interventions appeared to be moderately effective per the IRD ($ES = .61$) for improving social interactions for some participants but not all, indicating the possibility that a specific participant or a specific intervention feature influenced the strategy's effectiveness. Of the four studies in which the authors reported that the interventions were significantly effective, three appeared to have a very large effect sizes ($ES = 1.00$, $ES = .80$,

$ES = .87$) and one appeared to have a large effect size ($ES = .71$). There was considerable variation between the other eight studies' effect sizes notwithstanding the reported intervention efficacy. This may indicate that the intervention could be more or less efficacious under different conditions and that PND scores may be underestimating the effectiveness of the interventions.

The SMRS was developed to evaluate the rigor of intervention methodologies. The scientific rigor of published studies varies significantly and poorly-controlled studies are sometimes published due to interesting results that will encourage researchers to undertake better-controlled research (NAC, 2015). It is important that future research include objective and standardized quality indicators as a means to evaluate interventions. Although there are two studies in the current review that indicated initial evidence, the mean of the total SMRS scores for all studies indicated scientific rigor ($M = 3.41$). Studies showed moderate effectiveness according to the IRD metric scores, and methodological rigor was satisfactory for most of the studies. Although total SMRS scores of 3, 4, or 5 indicate the sufficient scientific rigor, there is a large gap between the scores of 3 and 5. Consequently, social story interventions related to social interaction still need additional and consistent research to support outcomes. The SMRS may be an effective way to structure and report results of the interventions.

Limitations

There are at least two limitations to this analysis that are important to consider in combination with these results. First, the application of rigorous selection criteria resulted in a small sample size of articles reviewed. Furthermore, analysis of some of the variables was based on an even smaller subsample of studies. Second, given that the results of this review were based on specific studies' results and from calculated IRD metric scores, the results of the review should be viewed with caution. The use of IRDs to measure treatment outcomes for single-subject studies is somewhat controversial because of the lack of statistical justification. Distributional properties of nonparametric methods are unknown,

so standard errors may not be justified (What Works Clearinghouse, 2010).

Directions for Future Research

A review of the current literature suggests that social stories are interventions that operate in multiple ways and produce varying results. There is no conclusive evidence to support a specific format or method of implementation for providing effective social stories either alone or with other components. The confounding of social story interventions with additional strategies is a problem in many existing studies (Kuoch & Mirenda, 2003). Although there are a number of issues that need to be addressed in future research, several emerge of particular importance. First, inadequate descriptions of participants make it difficult to determine whether participant-related variables moderate the effects of the interventions. Adequate participant descriptions, including information pertaining to the participants' cognitive abilities, severity in ASD, and language abilities, should be a component of future research for social story interventions.

Likewise, future researchers should employ NAC guidelines as quality indicators of their research to provide stronger intervention evaluation. Further, examining social validity is important; stakeholder concerns have not been fully explored in either the studies reviewed or previous research on social stories. Social validity refers to the social significance of the goals, the social appropriateness of the procedures, and the social importance of the effects (Wolf, 1978). One of the most objective forms of social validity is conducting comparisons with typical peers via direct observation (Ennis, Jolivette, Fredrick, & Alberto, 2013). Peer comparison could be used for all single-subject study designs since examining a typical peer's behavior can provide insight into the social acceptability of treatment (Ennis et al., 2013).

Conclusion

This review provides evidence that social story interventions can have a positive impact on the social interaction of students with ASD. Given the great variability in research environ-ments in terms of story format, implementation methods, and additional strategies, it is difficult to ascertain whether the social story or another component of the intervention is the critical reason for the decrease or increase in target behavior. Additional studies are still needed to examine the efficacy of social story interventions on social interactions for students with ASD.

References

Articles included in this review are marked with an asterisk.

Agosta, E., Graetz, J. E., Mastropieri, M. A., & Scruggs, T. E. (2004). Teacher-researcher partnerships to improve social behavior through Social Stories. *Intervention in School and Clinic, 39,* 276–287.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

Barnard-Brak, L., Ivey-Hatz, J., Ward, A. K., & Wei, T. (2014). Self- regulation and social interaction skills among children with autism across time. *Advances in Mental Health and Intellectual Disabilities, 8,* 271–279.

Baron-Cohen, S. (1995). *Mindblindness.* Cambridge, MA: MIT Press.

*Barry, L., & Burlew, S. (2004). Using Social stories to teach choice and play skills to children with autism. *Focus on Autism and Other Developmental Disabilities, 19,* 45–51.

Bellini, S., Peters, J., Benner, L., & Hopf, A. (2007). A Meta- Analysis of School-based social skills interventions for children with autism spectrum disorders. *Journal of Remedial and Special Education, 28,* 153–162.

Bledsoe, R., Myles, B. S., & Simpson, R. L. (2003). Use of Social Story™ intervention to improve mealtime skills of an adolescent with Asperger syndrome. *Autism, 7,* 289–295.

Boucenna, S., Narzisi, A., Tilmont, E., Muratori, F., Pioggia, G., Cohen, D., & Chetouani, M. (2014). Interactive technologies for autistic children. *Cognitive Computation, 6,* 145–166.

Crozier, S., & Tincani, M. J. (2005). Using a modified Social Story to decrease disruptive behavior of a child with autism. *Focus on Autism and Other Developmental Disabilities, 20,* 150–157.

*Crozier, S., & Tincani, M. J. (2007). Effects of Social Stories on prosocial behavior of preschool children with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 37,* 1803–1814.

*Delano, M., & Snell, M. E. (2006). The effects of Social Stories on the social engagement of chil-

dren with autism. *Journal of Positive Behavior Interventions, 8,* 29–42.

Dettmer, S., Simpson, R. L., Myles, B. S., & Ganz, J. B. (2000). The use of visual support to facilitate transitions of students with autism. *Focus of Autism and Other Developmental Disabilities, 15,* 163–169.

Ennis, R. P., Jolivette, K., Fredrick, L. D., & Alberto, P. A. (2013). Using comparison peers as an objective measure of social validity: Recommendations for researchers. *Focus on Autism and Other Developmental Disabilities, 28,* 195–201.

Garfield, J., Peterson, C., & Perry, T. (2001). Social cognition, language acquisition and the development of the theory of mind. *Mind & Language, 16,* 494–541.

Gray, C. (1995). Teaching children with autism to read social situations. In K. Quill (Ed.), *Teaching children with autism: Strategies to enhance communication and socialization* (pp. 219–242). Albany, NY: Delmar.

Gray, C. (1998). Social Stories and comic strip conversations with students with Asperger syndrome and high-functioning autism. In E. Schopler, G. B. Mesibov, & L. J. Kunce (Eds.), *Asperger Syndrome or high functioning autism?* (pp. 167–198). New York: Plenum.

Gray, C. (2010). *The new social story book.* Arlington, TX: Future Horizons.

Gray, C., & Garand, J. (1993). Social Stories: Improving responses of students with autism with accurate social information. *Focus on Autistic Behavior, 8,* 1–10.

Hobson, R. P. (1993). *Autism and the developmental of mind.* Hove, E. Sussex: Erlbaum.

Individuals with Disabilities Education Act, 20 U. S. C. 1400 (2004).

*Kagohara, D. M., Achmadi, D., Meer, L., Lancioni, G. E., O'Reilly, M. F., Lang, R., . . . Sigafoos, J. (2013). Teaching two students with Asperger syndrome to greet adults using social stories and video modeling. *Journal of Developmental and Physical Disabilities, 25,* 241–251.

Kokina, A., & Kern, L. (2010). Social story interventions for student with Autism Spectrum Disorders: A Meta-analysis *Journal of Autism and Developmental Disorders, 40,* 812–826.

Kuoch, H., & Mirenda, P. (2003). Social story interventions for young children with autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities, 18,* 219–228.

Leslie, A. (1987). Pretence and representations: The origins of "theory of mind." *Psychological Review. 94,* 412–426.

Locke, J., Ishijima, E. H., Kasari, C., & London, N. (2010). Loneliness, friendship quality and the social networks of adolescents with high-functioning autism in an inclusive school setting. *Journal of Research in Special Educational Needs, 10,* 74–81.

*Malmberg, D. B., Charlop, M. H., & Gershfeld, S. J. (2015). A two experiment treatment comparison study: Teaching social skills to children with Autism Spectrum Disorder. *Journal of Developmental and Physical Disabilities, 27,* 375–392.

Matson, L. J., Dempsey, T., & LoVullo, A. V. (2009). Characteristics of social skills for adult with intellectual disability, Autism and PDD-NOS. *Research in Autism Spectrum Disorders, 3,* 207–213.

Mayton, M. R., Menendez, A. L., Wheeler, J. J., Carter, S. L., & Chitiyo, M. (2013). An analysis of Social Stories research using an evidence-based practice model. *Journal of Research in Special Educational Needs, 13,* 208–217.

National Autism Center. (2015). *Findings and conclusions: National standards project Phase 2.* Randolph, MA: Author.

*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.

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

Reynhout, G., & Carter, M. (2006). Social stories for children with disabilities. *Journal of Autism and Developmental Disorders, 36,* 445–469.

Reynhout, G., & Carter, M. (2011). Evaluation of the efficacy of Social Stories using three single subject metrics. *Research in Autism Spectrum Disorder, 5,* 885–900.

Rohatgi, A. (2015). WebPlotDigitizer. http://arohatgi.info/WebPlotDigitizer

*Sansotti, 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.

*Sansotti, 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 disorder. *Journal of Positive Behavior Interventions, 10,* 162–178.

*Scattone, D. (2008). Enhancing the conversation skills of a boy with Asperger's disorder through social stories and video modeling. *Journal of Autism and Developmental Disabilities, 38,* 395–400.

*Scattone, D., Tingstrom, D. H., & Wilczynski, S. M. (2006). Increasing appropriate social interactions of children with autism spectrum disorders using social stories. *Focus on Autism and Other Developmental Disabilities, 21,* 211–222.

Scattone, D., Wilczynski, S. M., Edwards, R. P., & Rabian, B. (2002). Decreasing disruptive behaviors of children with autism using Social Stories. *Journal of Autism and Developmental Disorders, 32,* 535–543.

*Schneider, N., & Goldstein, H. (2010). Using So-

cial stories and visual schedules to improve socially appropriate behaviors in children with autism. *Journal of Positive Behavior Interventions, 12,* 149–160.

Test, D. W., Richter, S., Knight, V., & Spooner, F. (2011). A comprehensive review and meta-analysis of the social stories literature. *Focus on Autism and Other Developmental Disabilities, 26,* 49–62.

*Thiemann, K., & Goldstein, H. (2001). Social stories, written text cues, and video feedback: Effects on social communication of children with autism. *Journal of Applied Behavior Analysis, 24,* 425–446.

Trevarthen, C., Aitken, K., Papudi, D., & Roberts, J. (1996). *Children with autism: Diagnosis and intervention to meeting their needs.* London: Jessica Kingsley.

Vannest, K. J., Parker, R. I., & Groen, O. (2011). *Single case research: Web based calculators for SCR analysis.* (Version 1.0) [Web-based application]. College Station, TX: Texas A&M University. http://singlecaseresearch.org

Volkmar, F. R., Sparrow, S. S., Rende, R. D., & Cohen, D. J. (1989). Facial perception in autism. *Journal of Child Psychology and Psychiatry, 30,* 591–598.

What Works Clearinghouse. (2010). *Procedures and standards handbook 2.0.* Retrieved from http://ies.ed.gov/ncee/wwc/pdf/wwc_procedures_v2_standards_handbook.pdf

Wolf, M. M. (1978). Social validity: The case for subjective measurement. *Journal of Applied Behavior Analysis, 11,* 203–214.