

# Social Story™ Interventions for Students with Autism Spectrum Disorders: A Meta-Analysis

Anastasia Kokina · Lee Kern

Published online: 7 January 2010  
© Springer Science+Business Media, LLC 2010

**Abstract** A meta-analysis of single-subject research was conducted, examining the use of Social Stories™ and the role of a comprehensive set of moderator variables (intervention and participant characteristics) on intervention outcomes. While Social Stories had low to questionable overall effectiveness, they were more effective when addressing inappropriate behaviors than when teaching social skills. Social Stories also seemed to be associated with improved outcomes when used in general education settings and with target children as their own intervention agents. The role of other variables of interest, such as participants' age, diagnosis, and skill development, the format of Social Stories, the length of the intervention, and the use of assessment (e.g., comprehension checks) also was explored.

**Keywords** Autism spectrum disorders · Social Stories · Meta-analysis

## Introduction

In the context of the general call for evidence-based practices and the increase in the number of interventions for children and adolescents with autism spectrum disorders (ASD), the need to critically evaluate intervention effectiveness becomes more important than ever (Heflin and Simpson 1998). As many interventions emerge from clinical practice and require scientific validation, the dialogue between research and practice may become

complicated. Social Stories™ for children with ASD provide a good illustration of this scenario. While this intervention has a strong practical rationale and may be appealing to parents and practitioners, some researchers have argued that its scientific base is yet to be established (e.g., Sansosti et al. 2004).

Social Stories, first introduced in 1993 by educational consultant and former teacher Carol Gray, are primarily aimed at assisting individuals with ASD with their social difficulties (Gray 1998). Social dysfunction has been described as a universal, defining impairment in ASD (American Psychiatric Association 2000; Carter et al. 2005). Social characteristics in ASD may range from withdrawal or passive acceptance of others' initiations to high social motivation and frequent interactions, albeit not necessarily appropriate in nature. In general, however, observational studies indicate that students with ASD have lower social engagement and less frequently initiate and respond to initiations than their typical peers (Jackson et al. 2003; Jahr et al. 2007). Rules of the social world may be confusing and overwhelming even for high-functioning individuals with ASD (e.g., Grandin and Scariano 1986). As a result, social prognosis of people with ASD is often poor, including having experiences of loneliness, difficulty establishing and maintaining social relationships, and a range of mental health problems later in life (Bauminger et al. 2003; Howlin et al. 2004; Orsmond et al. 2004).

The primary goal of Social Stories is to address those debilitating difficulties. Social Stories are short stories written with the goal of objectively sharing important social information with individuals with ASD (Gray 1998, 2004). They explain difficult social situations and concepts in simple words, thereby providing students with “practical, tangible social information” (Gray 1998, p. 169) about a person, skill, event, concept, or situation. Gray

---

A. Kokina (✉) · L. Kern  
Special Education Program, College of Education, Lehigh University, 111 Research Drive, Bethlehem, PA 18015, USA  
e-mail: ank205@lehigh.edu

(2004) emphasized that Social Stories are in no way a tool for behavior change; rather, the premise of the intervention is that better social understanding will lead to improvements in behavior and social functioning.

To meet the defining criteria proposed by Gray (2004), a Social Story must include several types of sentences: (a) descriptive—factual statements used to describe the situation and people involved in it; (b) perspective—descriptions of the reactions, feelings, and responses of others; (c) directive—statements that identify an appropriate response and guide child's behavior; (d) cooperative—sentences to identify what others will do to assist; (e) affirmative—statements that enhance the meaning by expressing values or opinions common in a given culture; and (f) control—sentences written by the child to identify his/her personal strategies to recall and use information. Gray (2004) recommends using the ratio of one directive sentence to two or more sentences of the other types in every Social Story. This is important for students to have enough information and to avoid the Social Story becoming merely a list of things to do (Gray 1998). Social Stories are typically written from the first- or third-person perspective. It is recommended to avoid terms or statements that are “inflexible” as students with ASD may interpret them literally; instead, the terms “usually”, “sometimes”, and “probably” are used. Social Stories are presented in a written format, with or without illustrations, and are read either by adults or by the students themselves, just prior to a problematic situation (see Gray 1998). The Social Story may stay with a child as a permanent reminder, and may be used independently of adult prompts. Implementation is carefully monitored and students' progress is constantly evaluated. Additions or revisions to the Social Story are made based on the changes in situation, context, or student's behavior.

The theoretical rationale for Social Story interventions has been established in Gray's early writings (e.g., Gray 1998). According to one of the recognized cognitive theories of ASD, social difficulties of individuals with ASD may be linked to deficits in theory of mind (TOM) or the impaired ability to understand behaviors of others based on their beliefs, knowledge, desires, and feelings (e.g., Baron-Cohen et al. 2005). Difficulties associated with the impaired TOM may be addressed in Social Stories, which explain the difficult social concepts in simple words, and often include the description of views, perspectives, and feelings of other persons (i.e., perspective sentences). In another theory of ASD, social difficulties are linked to “weak central coherence” (WCC) or the impaired ability to derive generalized meaning from the context, while having a preference for detail-focused processing (Happé 2005). This means that in some situations individuals with ASD may pay attention to irrelevant details and fail to

understand the meaning of those situations. Many Social Stories are written to explain the meaning of problematic situations to the students and emphasize the relevant details, thereby addressing students' difficulties stemming from WCC. Finally, elements of Social Stories make them particularly appropriate for addressing several areas of relative strength and need in ASD: the need for predictability (American Psychiatric Association 2000), difficulty in acquiring long response chains (MacDuff et al. 1993), and preference for visually cued instruction (Quill 1997). Social Stories are often presented in a visual mode, with small fragments of information at a time, and are commonly read in advance of the targeted situation thereby increasing predictability. In incorporating those elements, Social Stories are similar to several well-researched interventions commonly used with students with ASD: task analysis, activity schedules, and priming. The aforementioned defining criteria described by Gray (2004), however, is what makes Social Stories distinct from other interventions.

In addition to the theoretical rationale, the relative ease of implementation makes Social Stories an attractive intervention option for practitioners and parents attempting to improve social outcomes of children with ASD. Social Stories are viewed by many teachers as a feasible and effective intervention (Smith 2001). The research findings, however, often suggest the opposite. Several descriptive reviews of the literature (e.g., Ali and Frederickson 2006; Nichols et al. 2005; Rust and Smith 2006; Sansosti et al. 2004) were published between 2004 and 2006 and synthesized almost identical pools of the studies. In the earliest of those reviews, Sansosti et al. (2004) criticized Social Story research on several methodological grounds, such as lack of experimental control, weak treatment effects, lack of maintenance and generalization data, and problems with the integrity of implementation. Several parallel reviews (e.g., Ali and Frederickson 2006; Rust and Smith 2006) recognized Social Stories as potentially effective, but similar to Sansosti et al., called for: (a) the improved methodological quality of future research, including more rigorous research designs and (b) further examination of the critical intervention variables that may moderate effectiveness (e.g., the role of the Social Story sentence ratio, participant characteristics).

The only quantitative review published to date, a meta-analysis by Reynhout and Carter (2006), provided a comprehensive coverage of the published and unpublished empirical studies (a total of 16 studies) conducted prior to December 2003. The descriptive synthesis was supplemented by a calculation of the percentage of non-overlapping data (PND) for studies using the single-subject designs and effect size calculations for group design studies. Results suggested low/questionable overall

effectiveness of Social Stories, with a wide range of individual outcomes. The review was mainly qualitative in nature, with the quantitative index of effectiveness (i.e., PND) used to examine only four intervention variables: “Social Story ratio”, comprehension checks, accessibility of Social Stories, and the use of additional strategies. Interestingly, Social Stories that deviated from the ratio (i.e., included more directive than descriptive sentences) seemed to produce better intervention outcomes than those that followed Gray’s criteria. Furthermore, studies that included comprehension checks yielded higher mean PND scores than those that did not assess comprehension. The differences between PND scores obtained for the other two variables (i.e., accessibility and the use of additional strategies) were minimal. Although the study by Reynhout and Carter provided several important insights into the use and effectiveness of Social Stories, it did not examine other variables (e.g., participant characteristics, goals of intervention) quantitatively. As a result, the role of those variables in moderating the effectiveness of Social Story interventions currently remains unexplored.

In summary, several common themes emerged from the existing syntheses of the literature on Social Stories. Most authors (e.g., Nichols et al. 2005; Reynhout and Carter 2006; Sansosti et al. 2004) agree that although Social Stories are a promising intervention, given the research to date, it is premature to conclude that they constitute an evidence-based strategy. Outcomes of many of the published studies need to be interpreted with caution due to their methodological limitations (e.g., the use of “treatment packages” and a lack of experimental control). Although limited, the available evidence suggests that the effectiveness of Social Stories is highly variable. That is, the intervention may be extremely effective with some students under certain conditions, but not others. The possible sources of this variability (e.g., student and intervention characteristics) warrant future investigation.

The present study differed from the previous reviews, including the meta-analysis by Reynhout and Carter (2006), in a number of important ways. First, the previous studies used loose inclusion criteria, including studies without methodological rigor or experimental control. This may be particularly problematic when conducting a meta-analysis of research, making firm conclusions about the effectiveness of intervention impossible. The current meta-analysis used rigorous selection criteria in order to control for the methodological quality of the included studies. Second, several previous reviews (e.g., Sansosti et al. 2004) failed to include unpublished dissertations in their analysis, including only published articles. This may lead to biased interpretation, due to the so-called “file drawer problem” (Rosenthal 1979) or the tendency to publish studies only with positive or significant outcomes. The

present review included both published studies and unpublished dissertations. Finally, most of the previous reviews were descriptive. The only study to use a quantitative effectiveness index (i.e., Reynhout and Carter 2006) examined the role of a limited set of intervention variables. The present study extended the findings of Reynhout and Carter (2006) by providing a detailed quantitative examination of a comprehensive set of variables that may moderate the effectiveness of Social Story interventions.

The present meta-analysis had three goals: (a) to examine the overall effectiveness of Social Story interventions, (b) to describe the ways in which Social Stories were used in research studies, and (c) to examine the role of a comprehensive set of moderator variables, including intervention and participant characteristics, on the effectiveness of Social Stories.

## Methods

### Study Identification and Selection

To identify studies for this meta-analysis, *PsycInfo* and *ERIC* electronic databases were searched using the combination of the terms “Social Story interventions”, “children” and “autism”. The search was restricted to English-language peer-reviewed documents published before April 2009. In addition, *ProQuest Dissertations* electronic database was searched to locate the unpublished dissertations. An ancestral search of the reference lists of the located empirical studies and reviews of the literature also was conducted to identify any additional references. Finally, several journals (i.e., *Focus on Autism and Other Developmental Disabilities*, *Journal of Positive Behavior Interventions*, *Journal of Autism and Developmental Disorders*) for 2008–2009 were hand searched. The search resulted in the identification of a total of 64 studies, including five case studies/concept papers, five research reviews (described previously), 18 doctoral dissertations, and 36 empirical studies.

To be included in the meta-analysis, the studies had to meet the following criteria: (a) used a single-subject design with demonstration of the experimental control (i.e., reversal design, multiple baseline across three or more legs) and graphically displayed baseline and intervention data to allow for calculation of the percentage of non-overlapping data (PND; Scruggs et al. 1987); (b) involved participants with a primary diagnosis of ASD made by an independent diagnostician, and (c) used Social Stories as a sole intervention. Excluded from the analyses were: (a) studies that did not contain quantitative data (e.g., qualitative case studies) or used group designs. Group design studies were excluded in order to use a uniform metric of treatment effectiveness (i.e., PND), because of

the difficulty combining the effect size measures of group design studies with PND, and because the number of group design studies was very small; (b) studies that used non-experimental AB designs or their variations (e.g., ABC); (c) studies that used Social Stories as part of a treatment package (e.g., Social Stories combined with verbal prompting, video modeling, role playing), unless the effectiveness of Social Stories alone was compared to that of a treatment package (e.g., Crozier and Tincani 2005); (d) studies that involved participants with disorders other than ASD (e.g., behavior disorders, intellectual disabilities); and (e) studies that had “floor” or “ceiling” effects in baseline, as evidenced by the presence of 0% data points for inappropriate behavior or 100% data points for appropriate behavior, making PND an inappropriate index of treatment effectiveness (Scruggs and Mastropieri 1998).

Of the 54 studies that contained participant data, a total of 36 were excluded for the following reasons: (a) the use of treatment packages without parceling out the effects of Social Stories alone ( $n = 18$ ); (b) the use of AB designs ( $n = 7$ ); (c) the use of group designs ( $n = 4$ ); note that two of those studies also used treatment packages; (d) the inclusion of participants whose primary diagnoses were not ASD ( $n = 3$ ); and (e) “floor” or “ceiling” effects in baseline data ( $n = 2$ ). Finally, two additional dissertations were excluded because the findings were published (i.e., Delano and Snell 2006; Scattone et al. 2002) and the studies were included in the final sample.

As a result, a total of 18 studies (15 articles and 3 dissertations) were included in the meta-analysis. Several datasets for individual participants within the included studies were subsequently excluded because of the presence of additional treatments or of “ceiling” or “floor” effects in the baselines (see Table 1 for a descriptive analysis of the included studies).

#### Data Coding and Analysis

To describe the use of Social Stories in research, each study was summarized and the information was coded into three broad categories: study methodology, intervention characteristics, and participant characteristics.

#### Study Methodology

The following study features were coded within this category: (a) experimental design (e.g., ABAB design, multiple baseline designs); (b) assessment of skill maintenance and generalization; (c) treatment fidelity (i.e., accuracy of implementation); (d) social validity (i.e., acceptability of intervention for stakeholders); and (e) inter-observer reliability information (i.e., consistency of measurement across the observers).

#### Characteristics of Intervention

This category contained the following codes: (a) goals of intervention, including reduction of problem behaviors (e.g., tantrums, crying, aggression), increase in social skills (e.g., appropriate play, social engagement, or initiations), acquisition of academic or functional skills (e.g., setting the table, counting, washing hands), and assisting students in transitions and novel situations; (b) intervention setting, including home, general education classroom, or special education classroom; (c) intervention agents (i.e., persons who administered the intervention), including teachers, researchers, target students, peers, or parents; (d) timing of intervention (i.e., whether Social Stories were read immediately prior to the targeted situation or some time in advance of it); (e) duration of intervention, including brief with 1–10 sessions, medium with 11–20 sessions, or long with 21–30 sessions; and (f) number of Social Stories per child (i.e., one or several). The following characteristics of Social Stories also were analyzed as part of this category: (a) format, including Social Stories with or without illustrations, presented via audio or computer equipment; (b) length of each Social Story (i.e., 1–5 sentences, 6–10 sentences, 11–15 sentences, or 16 and more sentences); (c) types of behaviors described, including singular simple behavior (e.g., hitting, use of appropriate tone of voice); complex social routine (e.g., greeting, having a conversation); or nonsocial routines (e.g., brushing the teeth, baking a cake); (d) use of functional behavior assessment (FBA) before intervention; and (e) use of comprehension checks post-treatment.

#### Participant Characteristics

This category included the following codes: (a) age (preschool—0–5 years old, elementary school—6–11 years old, middle school—12–14 years old, or high school—15–21 years old); (b) primary diagnosis (i.e., Asperger syndrome, pervasive developmental disorder—not otherwise specified [PDD-NOS], autism); (c) level of cognitive functioning (i.e., high or average—IQ scores at or above 70 on standardized assessments, or delayed—IQ scores below 70); (d) social skill development (i.e., high or average—the presence of high social motivation and use of appropriate social behaviors; limited—the presence of appropriate social skills, but a lack of their consistent use, appropriate responding, but a lack of appropriate initiations; or low—a lack of responsiveness, extreme social withdrawal, aversion to social contact); (e) communication skill development (high or average—scores of 70 and above on standardized tests of communication/language skills and/or the presence of strong receptive and expressive skills and the ability to speak in complete sentences; or limited—assessment scores below 70 on standardized tests of

**Table 1** Studies included in the meta-analysis

Study	N	Age, diagnosis	Target skills	Setting	Agent	Number of sessions	FBA	Comprehension	Fidelity	Social validity	Additional
Adams et al. (2004)	1	7, ASD	Frustration behaviors (crying, hitting)	Home	Parent	21–30	No	No	No	Yes	Two additional behaviors excluded due to floor effects
Bledsoe et al. (2003)	13	Asperger	Mealtime skills (food spills, napkin use)	Special education	Researcher	9	No	No	No	No	
Brownell (2002)	4	6–9, Autism	Echolalia, refusal to follow direction, use of loud voice	Special education	Researcher	10	No	No	No	No	
Crozier and Tincani (2005)	1	8, Autism	Talking out	Special education	Student	12	Yes	Yes	Yes	Yes	
Crozier and Tincani (2007)	3	3–5, ASD	Appropriate sitting; talking to peers; play	General education	Researcher	11–17	Yes	Yes	Yes	Yes	
Delano and Shell (2006)	2	6, Autism	Attention seeking comments, requests, responses	General education	Researcher	15	No	Yes	Yes	Yes	One participant excluded from the analysis due to the use of a treatment package
Demiri (2004) <sup>a</sup>	5	5–7, PDD-NOS	A total of 12 individual behaviors (e.g., sharing in play, eye contact, grabbing, nonfunctional toy use)	Home; general education	Teacher	4–29	No	Yes	Yes	Yes	Nine behaviors excluded due to floor/ceiling effects; two excluded due to treatment package
Dodd et al. (2008)	2	9–12, PDD-NOS	Excessive directions; compliment-giving	Home	Parent	3–4	No	Yes	Yes	Yes	
Graetz (2003) <sup>a</sup>	4	10–13, Autism	Use of high voice pitch; mouthing of fingers; repeating of “What”; hand-wringing	Special education	Teacher; student	5–21	No	No	No	Yes	One participant excluded due to the use of a treatment package
Ivey et al. (2004)	3	7, 5, 5, PDD-NOS	Participation in novel events	Home, community	Parent	8	No	No	Yes	Yes	
Keyworth (2004) <sup>a</sup>	3	12, 9, 7, ASD	Social initiations, responding, continuation, play	Special education	Researcher	15–21	No	Yes	Yes	Yes	Two behaviors excluded due to ceiling effects
Kuoch and Mirenda (2003)	2	5–6, ASD	Inappropriate behaviors during lunchtime or in games	General education	Teacher	5–8	No	No	Yes	No	One participant excluded due to floor effects
Lorimer et al. (2002)	1	5, ASD	Precursors to tantrum behaviors	Home	Teacher, parent	14	Yes	No	No	No	One behavior excluded due to floor effects
Ozdemir (2008)	3	7–9, ASD	Loud voice; chair tipping; cutting in lunch line	General education	Teacher, student	27	No	No	Yes	Yes	

**Table 1** continued

Study	N	Age, diagnosis	Target skills	Setting	Agent	Number of sessions	FBA	Comprehension	Fidelity	Social validity	Additional
Sansosti and Powell-Smith (2006)	3	9–11, Asperger	Treating peers with respect; conversation; joining in play	Special education, general education	Student	13–20	No	No	Yes	Yes	Yes
Scattone et al. (2002)	3	7–15, ASD	Tipping chair backward; staring; shouting	Special education	Student, teacher	7–18	No	Yes	Yes	Yes	Yes
Scattone et al. (2006)	3	8–13, ASD	Social initiation, response, comment, engagement	Special education, general education	Student, teacher	13–29	No	Yes	Yes	Yes	Yes
Schneider and Goldstein (2009)	3	5–10, Autism	On-task behaviors	Special education, general education	Researcher	12–33	No	Yes	No	No	No

<sup>a</sup> Unpublished dissertation

communication/language skills, the ability to use simple phrases or sentences to communicate, or no use of speech to communicate); and (f) reading skills (high or average—ability to independently read and comprehend complete sentences; limited—some reading skills, but a lack of fluency in reading complete sentences; or low—difficulty reading independently, pre-reading skills); and (g) the level of challenging behaviors—this category included, but was not limited to the targeted behaviors (i.e., high level—behaviors that may be harmful or potentially life-threatening, moderate level—behaviors that are disruptive and inappropriate, but not harmful or life threatening, or low level—behaviors that do not interfere with functioning, but may be distracting).

Several general rules were used to code participant characteristics. First, results of standardized assessments were used, if available. Scores within two standard deviations below the mean standard score of 100 were considered “average”. Second, if assessment results were unavailable, authors’ explicit qualitative descriptions were used (e.g., it was stated that participants had average to above average cognitive ability). Finally, if neither standard scores nor explicit descriptors were available, inferences regarding the participants’ skill levels could have been made (e.g., if a study alluded to the fact that the child was capable of having long conversations, it could have been interpreted as evidence of average communication skills; if a participant was at or above grade level academically, it could have been used as an indication that his cognitive abilities were average). In any given questionable case, a characteristic was not coded. Results of standardized assessments were available in many cases for cognitive functioning, but very rarely for communication skills.

To determine whether specific study characteristics were associated with higher intervention effectiveness, PND scores (Scruggs et al. 1987; Scruggs and Mastropieri 1998) were calculated. PND is a nonparametric approach to summarizing research, which determines the magnitude of behavior change from baseline to treatment phase by calculating the percentage of overlap between the data in those phases. Specifically, the number of intervention data points that exceed the highest or lowest baseline data point is divided by a total number of intervention data points, and the result is multiplied by 100%. PND is commonly used as the outcome metric in summaries of single-subject research as an alternative to the traditional visual methods of analysis. Its advantages include the ease of its calculation, agreement with visual analysis, and its applicability to any type of single-subject research design (see Parker et al. 2007). Scruggs and Mastropieri (1998, 2001) suggest that PND scores above 90 represent a highly effective intervention, scores from 70 to 90 represent effective

treatments, scores from 50 to 70 suggest outcomes that are questionable or low, and scores below 50 are ineffective.

Individual PND scores were calculated from each of the graphs provided in the studies and aggregated into summative scores for each participant and for each study. For ABAB reversal designs, following Scruggs and Mastropieri's (1998) recommendation, PND scores were calculated for each of the phases separately and then the two scores were aggregated (i.e., total number of non-overlapping data points was divided by total number of intervention data points). For multiple baseline designs, separate PNDs were calculated for each behavior, and then the individual PNDs were averaged to obtain the total score for the study. Mean PND scores for each study and for each participant were then grouped into coding categories for analysis. Study PND scores were used in the analysis of intervention characteristics (e.g., goals, settings, agents), while participant PNDs were used in the analysis of participant characteristics (e.g., age, diagnosis, communication skills). Whenever one study fit into several coding categories (e.g., within one study, Social Stories were implemented by teachers in one case and by parents in another, one participant was independent reader, while another child could not read), individual PND scores were used.

To summarize results of the intervention and to obtain the estimate of effectiveness for each of the study features, median PND scores were calculated. Median rather than mean was used in the summative analysis to control for the possible influence of outliers, as PND data are often not normally distributed. Statistical analyses were not conducted as part of this study, given the small sample size. As a result, all the analyses used in the study were descriptive.

The first author of the study served as the primary coder. Two other independent raters, doctoral students in Special Education, coded a random sample of 5 of the 18 (28%) studies and calculated PND scores. The index of inter-rater agreement was obtained by dividing the total number of agreements by the total number of agreements plus disagreements for each coded category, then multiplying the result by 100%. Total agreement for coding of the studies was 94% (range, 89–97%). Disagreements were mainly due to the difficulty in interpreting information on participant characteristics and were resolved through discussion. Agreement for PND calculations was 100%.

## Results

### Study Methodology

All the 41 published studies (both empirical and concept papers included in and excluded from the meta-analysis) were first examined to analyze the publication patterns.

These studies were published in a total of 18 journals. More than a half of the articles ( $n = 22$ , 54%) were published in one of the three journals: *Focus on Autism and Other Developmental Disabilities*, *Journal of Autism and Developmental Disorders*, or *Journal of Positive Behavior Interventions*. Research on Social Story interventions covers a relatively short time period. The recent years, however, witnessed a steady increase in the number of studies published on the topic. To illustrate, of the 41 articles on Social Stories published prior to spring of 2009, 33 were published between 2002 and 2009, while only eight studies were published between 1995 and 2001. All of the early studies were excluded after applying the criteria established for the present review. Of the 33 studies published between 2002 and the spring of 2009, 12 (36%) were published between 2002 and 2005, and 21 (64%) between 2006 and 2009.

The 18 studies included in the meta-analysis were published between 2002 and the spring of 2009 and involved a total of 47 participants. All of the studies employed single-subject designs; half of them ( $n = 9$ , 50%) used an ABAB reversal design or its variations, while another half utilized a multiple-baseline design. Half of the studies ( $n = 9$ , 50%) included in the review reported maintenance data, but only a few ( $n = 3$ , 17%) assessed generalization effects. All studies reported adequate inter-observer agreement. Finally, most of the studies ( $n = 13$ , 68%) reported information related to social validity of intervention, and the majority ( $n = 12$ , 67%) assessed treatment integrity.

### Overall Effectiveness and Characteristics of Intervention

A total median PND score of 62% and a mean score of 60% (range, 11–100%) were obtained for the intervention. Using the interpretation criteria suggested by Scruggs and Mastropieri (1998), this score places Social Stories in the low/questionable effectiveness category. Table 2 provides PND scores calculated for the variables examined in the current meta-analysis.

Differences in PND were examined relative to the specific type of behavior targeted for intervention. Eight studies (36%) targeted social behaviors, such as social engagement (e.g., Delano and Snell 2006); requests, responses, and initiations (e.g., Keyworth 2004; Sansosti and Powell-Smith 2006); appropriate play (e.g., Scattone et al. 2006); and nonverbal behaviors (Demiri 2004). Examples of challenging behaviors, targeted in 50% ( $n = 11$ ) of studies, included tantrum behaviors in homework preparation (Adams et al. 2004); screaming, talking out, or interrupting (Brownell 2002; Crozier and Tincani 2005; Lorimer et al. 2002); and inappropriately staring at

**Table 2** PND calculations for study variables

Study variables	Number of studies or participants (percentage)	<i>Mdn</i> PND (range)
<i>Intervention variables</i>		
Goal		
Improve appropriate social behaviors	8 (36%)	56 (19–92)
Reduce inappropriate behaviors	11 (50%)	87 (11–100)
Teach academic/functional skills	1 (5%)	22 (NA)
Assist in transitions, novel situations, reduce anxiety	2 (9%)	44 (38–50)
Setting		
Home	5 (23%)	30 (11–86)
General education	8 (36%)	73.5 (14–100)
Special education	9 (41%)	50 (14–100)
By whom read		
Teacher (general/special education)	7 (30%)	71 (18–100)
Researcher/data collector	6 (26%)	53 (17–92)
Target student	6 (26%)	95 (28–100)
Parents/caregiver	4 (18%)	55.5 (11–86)
When read		
Immediately before the target situation	13 (72%)	65 (11–100)
Not just before the target situation	5 (28%)	53 (30–88)
Duration of intervention		
Brief (1–10 sessions)	7 (28%)	71 (0–88)
Medium (11–20 sessions)	12 (48%)	66.5 (12–100)
Long (21–30 sessions)	6 (24%)	36.5 (5–100)
Number of Social Stories per participant		
One	14 (74%)	62 (11–100)
Several	5 (26%)	75 (25–92)
Social Story format		
Written story, without illustrations	5 (25%)	59 (11–88)
Written story and illustrations	14 (70%)	72 (17–100)
Musical format (i.e., song)	1 (5%)	95 (NA)
Length		
Brief (0–10 sentences)	10 (50%)	51 (15–100)
Long (11+ sentences)	10 (50%)	73 (7–100)
Types of behaviors addressed		
Singular behavior (social or nonsocial)	5 (25%)	87 (45–100)
Routine behavior (nonsocial)	6 (30%)	23.5 (0–88)
Social routines	9 (45%)	59 (23–100)
Use of functional assessment		
Yes	3 (17%)	86 (65–100)
No	15 (83%)	53 (11–100)
Use of comprehension checks		
Yes	9 (50%)	65 (25–100)
No	9 (50%)	53 (11–100)
<i>Participant characteristics</i>		
Grade/age		
Pre-K (0–5 years)	10 (21%)	50.5 (0–91)
Elementary (6–11 years)	28 (60%)	76.5 (11–100)
Secondary (12+ years)	9 (19%)	50 (0–94)



**Table 2** continued

Study variables	Number of studies or participants (percentage)	<i>Mdn</i> PND (range)
Primary diagnosis		
Asperger syndrome	4 (9%)	33.5 (15–95)
PDD-NOS	10 (21%)	38 (10–88)
Autism/ASD	33 (70%)	83 (0–100)
Cognitive functioning		
Cognitive abilities around average	26 (84%)	66 (11–100)
Cognitive delay	5 (16%)	71 (0–100)
Social skills		
Social skills limited	21 (64%)	50 (0–95)
Social skills low	12 (36%)	37.5 (0–100)
Communication skills		
Communication/language skills high or average	22 (59%)	83.5 (0–100)
Communication and language skills limited/low	15 (41%)	50 (0–100)
Reading skills		
High/average (fluent reader)	10 (26%)	79.5 (15–100)
Limited (some skills, but not fluent)	4 (10%)	84.5 (11–100)
Low (beginning/pre-reading skills)	25 (64%)	83 (0–100)
Level of challenging behaviors		
High—engages in dangerous behaviors	6 (13%)	45.5 (0–100)
Moderate—engages in disruptive behaviors	17 (38%)	71 (0–100)
Low—engages in tolerable behaviors	22 (49%)	74 (0–100)

females (Scattone et al. 2002). The less common targets of the intervention were academic or functional skills ( $n = 1$ , 5% of studies) and participation in novel events or transitions ( $n = 2$ , 9% of studies). Four studies addressed different types of behaviors and thus were included in several categories. Calculations of PND resulted in a median score of 56% (range, 19–92%) for social behaviors and 87% (range, 11–100%) for challenging behaviors. PND scores of 22 and 44% were obtained for studies targeting academic/functional skills and difficulties in transitions and novel situations, respectively. It should be noted that because only one study targeted academic or functional skills and two studies targeted participation in novel events or transitions, there are insufficient data to draw firm conclusions.

Most of the studies ( $n = 9$ , 41%) were conducted in special education settings (e.g., separate schools or self-contained classrooms). Other intervention settings included home ( $n = 5$ , 23%) and general education ( $n = 8$ , 36%). Four studies fit into several categories due to their use of different settings with different participants. Studies that were conducted in general education settings yielded a higher PND score (73.5%; range, 14–100%) than those conducted in special education settings (50%; range, 14–100%) or at home (30%; range, 11–86%). Social Story interventions were implemented by teachers ( $n = 7$ ; 30%),

researchers ( $n = 6$ , 26%), target students ( $n = 6$ ; 26%), and parents or caregivers ( $n = 4$ , 18%). Five studies used various intervention agents and were coded into several categories. Studies that involved target children as their own agents of intervention produced higher PND scores (95%; range, 28–100%) than those in which the Stories were read by teachers (71%; range, 18–100%), or parents (55.5%; range, 11–86%). The lowest PND scores were yielded by the studies that used researchers as intervention agents (53%; range, 17–92%).

In the majority of the studies ( $n = 13$ , 72%) Social Stories were read immediately prior to the situation in which the targeted behavior was most likely to occur (in a way similar to priming); the rest of the studies ( $n = 5$ , 28%) implemented the intervention in a delayed manner (i.e., read Social Stories some time before the targeted situation). Studies in which Social Stories were read just prior to the targeted situation had higher effectiveness scores (65%; range, 11–100%) than those in which Social Stories were not read just proximal to the situation (53%; range, 30–88%).

The moderating influence of intervention duration also was examined. Interventions ranged in length from 4 to over 30 sessions. Most studies described interventions that were of medium duration, or 11–20 sessions long ( $n = 12$ , 48%), while brief (1–10 sessions;  $n = 7$ , 28%) and long

interventions (over 20 sessions;  $n = 6$ , 24%) were less common. Several studies that used both brief and long interventions were included in more than one category. Brief interventions (i.e., 0–10 sessions) yielded higher scores (PND = 71%; range, 0–88%) than medium (PND = 66.5%; range, 12–100%) or long interventions (PND = 36.5%; range, 5–100%). To examine the possible effects of treatment intensity further, the analysis focused on the number of Social Stories each participant was exposed to as part of the intervention. Most studies ( $n = 14$ , 74%) used just one Social Story per participant at a given time. Five studies (26%) used several Stories per participant. The PND scores obtained for studies that used several Social Stories per child were higher (PND = 75%; range, 25–92%) than those that used just one Story (PND = 62%; range, 11–100%).

In addition, variables describing Social Story construction were examined. Social Stories were presented in the following formats: (a) written with illustrations (e.g., drawings or actual photographs;  $n = 14$ , 70%); (b) written without illustrations ( $n = 5$ ; 25%); and (c) musical format (i.e., song;  $n = 1$ , 5%). One study used both illustrated and written formats. Higher effectiveness (PND = 72%; range, 17–100%) was associated with the use of illustrated Stories than those without any illustrations (PND = 59%; range, 11–88%). Social Stories ranged in length from 6 to over 30 sentences. The same number of studies contained Social Stories that were brief, with less than 10 sentences ( $n = 10$ , 50%) and long, with 10 or more sentences ( $n = 10$ , 50%). Three studies used both brief and long Stories, depending on the participants' needs. Studies that used long Stories produced higher PND scores (PND = 73%; range, 7–100%) than those that used brief Stories (PND = 51%; range, 15–100%).

The content of each of the Social Stories was further examined in relation to the quality and complexity of behaviors described in them. This category was included in addition to the category of dependent variables because it seemed important to explore the additional aspects of behaviors addressed by the intervention. Most Social Stories ( $n = 9$ , 45%) included descriptions of social routines (e.g., playing with peers, having a conversation). Nonsocial routines (e.g., eating lunch) were described in about the same number of studies ( $n = 6$ , 30%) as simple singular behaviors (e.g., use of inappropriately loud voice, tipping chair backwards;  $n = 5$ , 25%). Two studies were coded in different categories, as they targeted different types of behaviors. Studies using Social Stories that addressed simple behaviors yielded higher scores (PND = 87%; range, 45–100%) than those that targeted social routines (PND = 59%; range, 23–100%), or nonsocial routines (PND = 23.5%; range, 0–88%).

Only a small number of studies ( $n = 3$ , 17%) used FBA information to guide construction and implementation of Social Stories. Several additional studies (e.g., Adams et al. 2004) reported conducting the functional assessment interviews, but did not report results of the assessment, and were coded as not having incorporated the FBA information. A total of 15 studies (83%) were coded as not using the FBA information. Studies that used FBA yielded higher effect sizes (PND = 86%; range, 65–100%) than studies that did not use it (PND = 53%; range, 11–100%). These findings should be interpreted with caution, however, given the small number of studies with FBA information. Finally, the same number of studies ( $n = 9$ , 50% each) did and did not include comprehension checks to assess participants' understanding of Social Stories. Studies that included assessment of comprehension yielded higher PND scores (65%; range, 25–100%) than those that did not (53%; range, 11–100%).

### Participant Characteristics

Individual participants' PND scores were calculated and then aggregated by participant characteristics. Participants' ages ranged between 3 and 15 years. Most of the study participants were younger children: preschool ( $n = 10$ , 21%) or elementary school ( $n = 28$ , 60%). Because only one participant was a high school student, the middle and high school grades were combined into the category "secondary school" ( $n = 9$ , 19%). Higher effectiveness scores were obtained for elementary school group (PND = 76.5%; range, 11–100%) than for preschool (PND = 50.5%; range, 0–91%), or secondary-school group (PND = 50%; range, 0–94%). Given that the elementary-aged participants constituted by far the largest group, the results should be interpreted with caution. Most of the study participants were diagnosed with autism ( $n = 33$ , 70%), Asperger syndrome ( $n = 4$ , 9%), or PDD-NOS ( $n = 10$ , 21%). None of the studies involved participants from other autism spectrum categories (e.g., Rett's syndrome, childhood disintegrative disorder). Higher effectiveness scores were obtained for the autism group (PND = 83%; range, 0–100%) than for participants with Asperger syndrome (PND = 33.5%; range, 15–95%), followed by students with PDD-NOS (PND = 38%; range, 10–88%).

Cognitive ability of 31 participants was coded. The majority of participants ( $n = 26$ , 84%) had high or average cognitive ability, whereas a smaller group ( $n = 5$ , 16%) had cognitive delay. Lower PND scores were obtained for the group with high to average cognitive ability (PND = 66%; range, 11–100%) than for the group of students who had cognitive delay (PND = 71%; range,

0–100%). However, the cell sizes were again highly unequal so results should be interpreted with caution.

Social skills of most of the participants, typical for many students with ASD, were limited ( $n = 21$ , 64%) or low ( $n = 12$ , 36%). None of the participants had high or average social skills, so this category was dropped. Students who had limited social skills seemed to benefit from Social Stories to a greater extent (PND = 50%; range, 0–95%) than those who had low social skills (PND = 37.5%; range, 0–100%). With regard to communication skills, most participants ( $n = 22$ , 59%) had high or average receptive and expressive language abilities. The limited skill category included 15 participants (41%). Higher effectiveness scores (PND = 83.5%; range, 0–100%) were obtained for participants with high or average communication skills than for those with limited skills (PND = 50%; range, 0–100%).

As reading skills may be pre-requisite for success of Social Story interventions, their relation to treatment outcomes was examined. Children with low (i.e., pre-reading) skills constituted the largest group ( $n = 25$ , 64%), followed by the group of children with high/average skills ( $n = 10$ , 26%), and those with limited skills ( $n = 4$ , 10%). Similar PND scores were obtained for the groups of children who had limited reading skills (PND = 84.5%; range, 11–100%), those who had low skills (PND = 83%; range, 0–100%), followed by the group of average or advanced readers (PND = 79.5%; range, 15–100%). Finally, the impact of challenging behaviors was examined. Children with low levels of challenging behaviors constituted the largest group ( $n = 22$ , 49%), whereas smaller groups of participants had moderate ( $n = 17$ , 38%) or high ( $n = 6$ , 13%) levels of challenging behaviors. The highest PND scores were obtained for the group with low levels of challenging behavior (PND = 74%; range, 0–100%), followed children with moderate (PND = 71%; range, 0–100%) or high-level behaviors (PND = 45.5%; range, 0–100%).

## Discussion

### Overall Effectiveness of Social Stories

Results of this investigation confirmed previous findings regarding the questionable effectiveness of Social Story interventions for students with ASD. Specifically, the total average intervention PND score in this study (60%; range, 11–100%) was somewhat higher than the score obtained in the meta-analysis by Reynhout and Carter (2006) (i.e., total PND = 51%; range, 20–95%), but still fell below the cut-off PND score of 70 suggested for effective interventions (Scruggs and Mastropieri 1998). These findings are

noteworthy and somewhat surprising given the accumulating number and the improved quality of the recent studies. Similarity of the results obtained in this study and the previous meta-analysis is also surprising given the following differences between the studies. First, the present investigation included a more recent sample of studies, with an overlap between the two meta-analyses of five studies only. Unlike in the previous review, a major effort was made to control for the methodological quality of research by excluding studies that used non-experimental designs, treatment packages, and studies with “floor” or “ceiling” effects in baseline, which may result in the artificially low PND scores. Yet even with the exclusion of those studies, the outcomes of the intervention were questionable. Both the present study and the meta-analysis by Reynhout and Carter (2006), however, revealed the extreme variability of the individual outcomes, manifested in a wide range of PND scores across the studies. Most of the individual participants’ PND scores fell into two effectiveness categories (Scruggs and Mastropieri 1998): “highly effective/effective” with PND scores over 70% ( $n = 24$ ) and “ineffective” with PND scores below 50% ( $n = 21$ ), forming the almost equally sized groups. Only two individual scores fell into the “questionable effectiveness” range. Therefore, the use of the intervention and the impact of the possible moderator variables were examined further.

### Intervention Characteristics

Results of this meta-analysis indicated that while Social Stories have been used to address a number of behaviors, the two main intervention goals were reduction of inappropriate behaviors and improvements in social skills. Other possible applications suggested by Gray (1998, 2004), such as teaching academic skills, assisting students in novel events, and acknowledging students’ achievements, remain largely unexplored in the research literature. For example, only two studies (Ivey et al. 2004; Schneider and Goldstein 2009) examined the use of Social Stories to address difficulties in novel situations and transitions, characteristic of students with ASD (American Psychiatric Association 2000). Reynhout and Carter’s (2006) meta-analysis resulted in similar conclusions, pointing to a limited use of Social Stories. Furthermore, the intervention seemed to be substantially more effective when used to target behavior reduction than to teach the appropriate social skills. One possible explanation is that social behaviors, situations, and concepts are more abstract and complex and thus more difficult for students to understand. Consistent with the “complexity” hypothesis, the intervention had higher effectiveness with Social Stories describing simple singular behaviors rather than complex social or nonsocial routines.

Alternatively, it may be that Social Stories promote an understanding of social concepts or situations (consistent with Gray 2004), and that improved social understanding results in reduced challenging behaviors, but not in improved social skills. That is, a student may understand a social situation or concept, but may lack social skills to apply this knowledge. If so, when designing a Social Story intervention, is it important to consider the skills needed to successfully achieve the intervention goals. If a student is lacking the pre-requisite social skills, the use of direct teaching of those skills may be required to supplement Social Stories. Gray's (1998) writings seem to support this notion. Specifically, she suggests that social understanding may be an important "*pre-requisite* [italics added] component to teaching social skills" (p.169). In general, teaching social skills to students with ASD is a challenging task, which must involve consideration of the multiple factors, including peers' responsiveness (e.g., Scattone et al. 2006). Careful planning of the intervention, including the identification of pre-requisite skills and environmental supports, is essential for its success.

Most studies were conducted in the self-contained settings. This seems unfortunate as Social Stories seem to be a good fit for the general education environments due to their ease of implementation and a relative unobtrusiveness. Supporting this notion, results of this study indicated that Social Stories implemented in the general education settings produced substantially larger effects on students' behaviors than those implemented in the self-contained settings. Furthermore, studies that used target children as their own intervention agents were substantially more effective than those that were run by adults (i.e., teachers, researchers, or parents). This suggests, on the one hand, that encouraging self-determination and independence in children and implementing Social Stories in general education settings may produce greater intervention benefits. At the same time, it is possible that children who are capable of reading and monitoring their own intervention are likely to be more successful. Similarly, students with higher level of skill development may be more likely to be included in general education. To illustrate, of the 26 students whose cognitive level was coded as high or average, 11 received the intervention in general education setting, 4 in special education, and 11 at home. At the same time, four out of five students with delayed cognitive skills were in special education, whereas one was in general education. Therefore, students' level of skill development may mediate the relationship between other variables (e.g., setting, agent of intervention) and treatment outcomes (see the discussion of participant characteristics below). Further, implementation by natural intervention agents (i.e., teachers or students) resulted in more pronounced intervention effects than implementation by researchers. The

relatively low PND scores obtained for the studies involving parents as intervention agents were possibly due to lower treatment fidelity. Similarly, low treatment effects were produced by the interventions conducted in the home settings. Thus, while it is important to use natural implementation agents in typical settings, treatment fidelity must be carefully monitored.

As noted, Social Stories include elements of several well-established interventions for students with ASD, such as priming or introducing the child to activities in analogue situations just prior to the actual situations (e.g., Koegel et al. 2003). The fact that intervention effectiveness was higher for Social Stories, which were read immediately before the target situation, in a manner similar to priming, is therefore unsurprising. This finding may be clarified in future experimental research by manipulating the amount of time passed between the reading of Social Stories and the problematic situation.

Two intervention characteristics, intervention length and the number of Social Stories per child, were included to examine the effects of treatment intensity. Brief interventions (i.e., 1–10 sessions) were associated with higher treatment effectiveness than medium (i.e., 11–20 sessions) or long (over 20 sessions) interventions. It is possible that the intervention effects may wear off as a result of longer intervention duration. In many cases, Social Stories seemed to produce immediate changes in the levels of targeted behaviors (e.g., Ozdemir 2008; Lorimer et al. 2002). At the same time, while most studies used just one Social Story per participant, the few studies that used several Social Stories per child (e.g., Delano and Snell 2006; Dodd et al. 2008; Lorimer et al. 2002) seemed to produce higher effects on the students' behavior. It is therefore possible that higher treatment intensity is associated with improved participant outcomes. Another hypothesis is that as a result of the exposure to several Social Stories children with ASD gain experience and become more proficient using them. If experience is indeed a moderating variable, initial instruction in the use of Social Stories may be required. This assertion finds some support in the fact that verbally reminding children to use the targeted behaviors resulted in improved behavior relative to the unprompted use in a number of studies (e.g., Crozier and Tincani 2005, 2007). The role of experience should be examined in future research. Future studies also should examine the number of daily readings of each Social Story as another indicator of treatment intensity.

Research studies included in the review implemented Social Stories in two formats, written and written with illustrations. It has been suggested that children with ASD respond well to visually cued methods of instruction (Quill 1997), so it is not surprising that many researchers used illustrations to enhance the content of Social Stories.

Moreover, Social Stories that used illustrations were more effective than those which used the written text only. Interestingly, while Gray (2004) indicated that Social Stories are used by the practitioners and parents in a variety of formats, such as Power Point presentations, Stories embroidered on a quilt, and Stories acted out by puppets, none of the experimental studies have employed such alternate methods of intervention delivery. It should be mentioned, however, that a few studies that were excluded from this meta-analysis used other methods of Social Story delivery, such as video-modeled and computer stories (e.g., Hagiwara and Myles 1999; Sansosti and Powell-Smith 2008). Only one included study (Brownell 2002) employed an unusual musical format of Social Story delivery. The Brownell (2002) study also compared the effects of musical Social Stories to the traditional teacher-read method, with the former yielding somewhat superior outcomes (i.e., PND = 95% for musical format and 87% for written/read format). While format and delivery methods should be individually determined by the needs and interests of students, the role of Social Story format certainly deserves to be further examined in research.

Although Carol Gray does not formally recommend the use of FBA as part of Social Story construction and implementation, the assessment process she described (1998, 2004) is in many ways similar to the process of FBA. The intervention begins with a comprehensive assessment of the individual, contexts, situations, and the underlying motivations for behavior. While there were only three studies in this review that used the FBA to guide Social Story interventions, those studies obtained substantially higher effectiveness scores than the investigations that did not use the FBA. Although preliminary, those results suggest that FBA has a potential in being able to inform Social Story interventions, and may be associated with favorable outcomes.

Further, comprehension checks may be an important part of Social Story implementation. Indeed, it seems reasonable to check children's level of understanding of the concepts and situations described in the Social Stories. In this meta-analysis, lower PND scores were obtained for the studies that did not involve comprehension checks, implying that lower treatment outcomes could be due to a lack of participants' understanding of Social Stories. Similar results were obtained in the previous meta-analysis (Reynhout and Carter 2006); therefore, it is recommended that professionals and parents conduct at least brief comprehension checks when using Social Stories.

#### Participant Characteristics

Although Gray (1998) initially described Social Stories as a method to assist high-functioning students with ASD, she

later revised her recommendations, suggesting that the intervention may also be appropriate for students with a broader range of abilities, including low-functioning individuals. Analysis of the moderating influence of participants' cognitive skills provided some support to this proposition. Specifically, effects of the intervention seemed to be somewhat higher for participants with lower cognitive ability than for students with high or average intelligence. Interestingly, a different pattern was evident with regard to communication skills, in that higher levels of communication skills were associated with substantially higher effectiveness. Given that Social Stories is a language-based intervention, higher verbal ability may be required for it to be successful. Similarly, better social skills seemed to be associated with more favorable intervention outcomes.

Findings related to the moderating role of participants' skill development should be interpreted with caution, however, given that a very limited proportion of students in the present review were lower functioning (e.g., 16% in the delayed cognitive skills category as opposed to 84% in the high or average skills group). In general, the majority of participants in the included studies had the following profile: elementary-grade students with high or average cognitive abilities, strong communication and relatively strong social skills, pre-reading skills, and low to moderate levels of inappropriate behaviors. This finding is striking given that the intervention is often thought of as suitable to support diverse groups of students with ASD. However, evidence with regard to the effectiveness of Social Stories with lower functioning students is currently limited. In light of these results, caution should be used when planning Social Story interventions to assist students with ASD who have more significant cognitive, social, or language delays. Clearly, additional studies involving students with lower levels of skill development need to be conducted. A similar theme emerged from the study by Reynhout and Carter (2006), who stated the need to explore the use of Social Stories with individuals with ASD who have significant intellectual disabilities.

Two additional findings of this study related to participant characteristics warrant further clarification. First, the finding that treatment effects were higher for the autism group than for participants with Asperger syndrome or PDD-NOS should be viewed with caution given that the autism group was the largest. In addition, most participants with autism, similar to students with Asperger syndrome, were verbal and had average cognitive ability; hence, the possible overlap in characteristics between the two groups. Second, somewhat contrary to the initial expectations, Social Stories were slightly less effective when used with proficient readers than with students who had limited or low skills. Conversely, effectiveness of Social Stories was very

similar for groups of students with limited and low reading skills. Those findings may imply that, given modifications to the format and delivery, Social Stories may be appropriate for students with varying reading skill levels.

In summary, a combination of the following factors seemed to be associated with higher effectiveness of Social Story interventions: (a) targeting reductions in inappropriate behaviors, (b) implementation in the general education setting, (c) the use of target children as their own intervention agents, (d) Social Stories read immediately prior to the targeted situation, (e) Social Stories describing simple singular behaviors rather than complex “chains” of behaviors, (f) brief duration of intervention, (g) the use of functional assessment to inform the intervention, (h) the use of comprehension checks, (i) involving elementary-aged participants with higher levels of communication and social skills, and low or moderate levels of challenging behaviors. Whether one or several of these variables is determinant for intervention outcomes should be clarified in the future experimental research.

The results described above should be viewed as preliminary due to several limitations of this meta-analysis. First, application of the rigorous selection criteria resulted in a very small sample size. Furthermore, because many studies failed to provide information about the variables of interest (e.g., participant characteristics), analysis of some of the variables was based on even a smaller subsample of studies. Therefore, results do not imply that Social Stories are ineffective or should not be used under a set of circumstances different from those highlighted in this review. Finally, the analyses used in this review were descriptive, as there was not enough power for statistical procedures (e.g., nonparametric tests). Interpretation of the results was based on the analysis of differences between the PND scores; therefore, the interpretation of the magnitude of differences was somewhat arbitrary.

There are several recommendations for researchers interested in further examining Social Story interventions. First, although the methodological quality of Social Story research seems to have improved (e.g., most of the earlier studies were excluded from the meta-analysis, while the more recent studies met the inclusion criteria), additional methodologically robust investigations are needed. Studies should include data on generalization and maintenance of skills, social acceptability of the intervention, and treatment fidelity. Investigations that parcel out the effects of Social Stories from other methods also would be timely. Most of the studies of Social Stories have used single-subject research designs that seem appropriate given a highly individualized nature of the intervention. However, with only four studies (one published) using the group design methodology, additional group studies of Social Stories are needed. Second, additional applications of Social Stories

need to be examined. As this review suggests, the intervention has never been used to address academic skills. Two studies only (i.e., Ivey et al. 2004; Schneider and Goldstein 2009) used Social Stories to address students' difficulties in novel events and transitions. Furthermore, the use of Social Stories with additional student populations and participant groups (e.g., students with disabilities other than ASD, students with lower level of skill development, secondary school students) also needs to be explored. Comprehensive descriptions of participant characteristics should be included to provide profiles of “responders” and “non-responders” to intervention. Finally, while this meta-analysis offers preliminary findings, additional experimental studies are needed that would explore the critical variables associated with intervention effectiveness.

## References

- \*Adams, L., Gouvousis, A., VanLue, M., & Waldron, C. (2004). Social story intervention: Improving communication skills in a child with an autism spectrum disorder. *Focus on Autism and Other Developmental Disabilities, 19*, 87–94.
- Ali, S., & Frederickson, N. (2006). Investigating the evidence base of Social Stories. *Educational Psychology in Practice, 22*, 355–377.
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Association. (text revision).
- Baron-Cohen, S., Wheelwright, S., Lawson, J., Griffin, R., Ashwin, C., Billington, J., et al. (2005). Empathizing and systemizing in autism spectrum conditions. In F. R. Volkmar, R. Paul, A. Klin, & D. Cohen (Eds.), *Handbook of autism and pervasive developmental disorders* (pp. 628–639). Hoboken, NJ: Wiley.
- Bauminger, N., Shulman, C., & Agam, G. (2003). Peer interaction and loneliness in high-functioning children with autism. *Journal of Autism and Developmental Disorders, 33*, 489–507.
- \*Bledsoe, R., Myles, B. S., & Simpson, R. L. (2003). Use of a Social Story intervention to improve mealtime skills of an adolescent with Asperger syndrome. *Autism, 7*, 289–295.
- \*Brownell, M. D. (2002). Musically adapted Social Stories to modify behaviors in students with autism: Four case studies. *Journal of Music Therapy, 39*, 117–144.
- Carter, A. S., Ornstein Davis, N., Klin, A., & Volkmar, F. R. (2005). Social development in autism. In F. R. Volkmar, R. Paul, A. Klin, & D. Cohen (Eds.), *Handbook of autism and pervasive developmental disorders* (pp. 312–334). Hoboken, NJ: Wiley.
- \*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 behaviors of preschool children with autism spectrum disorders. *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 children with autism. *Journal of Positive Behavior Interventions, 8*, 29–42.
- \*Demiri, V. (2004). *Teaching social skills to children with autism using Social Stories: An empirical study*. Doctoral dissertation, Hofstra University. DAI-B, 65 (05), 2619.

- \*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.
- \*Graetz, J. E. J. R. (2003). Promoting social behavior for adolescents with autism with Social Stories. Ph.D. dissertation, George Mason University, United States—Virginia. Retrieved April 19, 2008, from ProQuest Digital Dissertations database (Publication No. AAT 3079339).
- Grandin, T., & Scariano, M. M. (1986). *Emergence: Labeled autistic*. Novato, CA: Arena.
- 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. (2004). Social Stories 10.0: The new defining criteria. *Jenison Autism Journal, 15*, 1–21.
- Hagiwara, T., & Myles, B. S. (1999). A multimedia social story intervention: Teaching skills to children with autism. *Focus on Autism and Other Developmental Disabilities, 14*, 82–95.
- Happé, F. (2005). The weak central coherence account of autism. In F. R. Volkmar, R. Paul, A. Klin, & D. Cohen (Eds.), *Handbook of autism and pervasive developmental disorders* (pp. 640–649). Hoboken, NJ: Wiley.
- Heflin, L. J., & Simpson, R. L. (1998). Interventions for children and youth with autism: Prudent choices in a world of exaggerated claims and empty promises. Part I: Intervention and treatment option review. *Focus on Autism and Other Developmental Disabilities, 13*, 194–211.
- Howlin, P., Goode, S., Hutton, J., & Rutter, M. (2004). Adult outcomes for children with autism. *Journal of Child Psychology and Psychiatry, 45*, 212–229.
- \*Ivey, M. L., Heflin, L. J., & Alberto, P. (2004). The use of Social Stories to promote independent behaviors in novel events for children with PDD-NOS. *Focus on Autism and Other Developmental Disabilities, 19*, 164–176.
- Jackson, C. T., Fein, D., Wolf, J., Jones, G., Hauck, M., Waterhouse, L., et al. (2003). Responses and sustained interactions in children with mental retardation and autism. *Journal of Autism and Developmental Disorders, 33*, 115–121.
- Jahr, E., Eikeseth, S., Eldevik, A., & Aase, H. (2007). Frequency and latency of social interaction in an inclusive kindergarten setting. *Autism, 11*, 349–363.
- \*Keyworth, P. L. W. (2004). The effects of Social Stories on the social interaction of students with autism. Ph.D. dissertation, The University of Iowa, United States—Iowa. Retrieved April 19, 2008, from ProQuest Digital Dissertations database (Publication No. AAT 3157989).
- Koegel, L. K., Koegel, R. L., Frea, W., & Green-Hopkins, I. (2003). Priming as a method of coordinating educational services for students with autism. *Language, Speech & Hearing Services in Schools, 34*, 228–235.
- \*Kuoch, H., & Mirenda, P. (2003). Social Story interventions for young children with autism spectrum disorder. *Focus on Autism and Other Developmental Disabilities, 18*, 219–227.
- \*Lorimer, P. A., Simpson, R. L., Myles, B. S., & Ganz, J. B. (2002). The use of Social Stories as a preventative behavioral intervention in a home setting with a child with autism. *Journal of Positive Behavior Interventions, 4*, 53–60.
- MacDuff, G. S., Krantz, P. J., & McClannahan, L. E. (1993). Teaching children with autism to use photographic activity schedules: Maintenance and generalization of complex response chains. *Journal of Applied Behavior Analysis, 26*, 89–97.
- Nichols, S. L., Hupp, S. D. A., Jewell, J. D., & Ziegler, C. S. (2005). Review of social story interventions for children diagnosed with autism spectrum disorders. *Journal of Evidence-Based Practices for Schools, 6*, 90–120.
- Orsmond, G. I., Krauss, M. W., & Seltzer, M. M. (2004). Peer relationships and social and recreational activities among adolescents and adults with autism. *Journal of Autism and Developmental Disorders, 34*, 245–256.
- \*Ozdemir, S. (2008). The effectiveness of Social Stories on decreasing disruptive behaviors of children with autism: Three case studies. *Journal of Autism and Developmental Disorders, 38*, 1689–1696.
- Parker, R. I., Hagan-Burke, S., & Vannest, K. (2007). Percentage of all non-overlapping data (PAND): An alternative to PND. *The Journal of Special Education, 40*, 194–204.
- Quill, K. A. (1997). Instructional considerations for young children with autism: The rationale for visually cued instruction. *Journal of Autism and Developmental Disorders, 27*, 697–713.
- Reynhout, G., & Carter, M. (2006). Social Stories for children with disabilities. *Journal of Autism and Developmental Disorders, 36*, 445–469.
- Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin, 86*, 638–641.
- Rust, J., & Smith, A. (2006). How should the effectiveness of Social Stories to modify the behavior of children of the autistic spectrum be tested? *Autism, 10*, 125–138.
- \*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.
- 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., & Kincaid, D. (2004). A research synthesis of social story interventions for children with autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities, 19*, 194–204.
- \*Scattone, D., Tingstrom, D. H., & Wilczynski, S. M. (2006). Increasing appropriate social interaction 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. (2009). Using Social Stories and visual schedules to improve socially appropriate behaviors in children with autism. *Journal of Positive Behavior Interventions, 11*, 1–12.
- Scruggs, T. E., & Mastropieri, M. A. (1998). Synthesizing single subject research: Issues and applications. *Behavior Modification, 22*, 221–242.
- Scruggs, T. E., & Mastropieri, M. A. (2001). How to summarize single-participant research: Ideas and applications. *Exceptionality, 9*, 227–244.
- Scruggs, T. E., Mastropieri, M. A., & Casto, G. (1987). The quantitative synthesis of single-subject research: Methodology and validation. *Remedial and Special Education, 8*, 24–32.
- Smith, C. (2001). Using Social Stories to enhance behavior in children with autistic spectrum difficulties. *Educational Psychology in Practice, 17*, 337–345.

References marked with an asterisk indicate studies included in the meta-analysis.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.