

Social Stories™ for Children with Disabilities

Georgina Reynhout^{1,2} and Mark Carter¹

A review of the empirical research literature on Social Stories™ is presented, including a descriptive review and single-subject meta-analysis of appropriate studies. Examination of data suggests the effects of Social Stories™ are highly variable. Interpretations of extant studies are frequently confounded by inadequate participant description and the use of Social Stories™ in combination with other interventions. It is unclear whether particular components of Social Stories™ are central to their efficacy. Data on maintenance and generalization are also limited. Social Stories™ stand as a promising intervention, being relatively straightforward and efficient to implement with application to a wide range of behaviors. Further research is needed to determine the exact nature of their contribution and the components critical to their efficacy.

KEY WORDS: Social Stories™; autism; Asperger's syndrome; behavior modification; social skills; special education.

INTRODUCTION

In describing Social Stories™, Gray (2003) has stated that:

A social story is a *process* that results in a *product* for a person with autism spectrum disorder (ASD). First, as a process, a social story requires consideration of—and respect for—the person with ASD. As a product, a social story is a short story—defined by specific characteristics—that describes a situation, concept, or social skill using a format that is meaningful for people with ASD. The result is often renewed sensitivity of others to the experience of the person with ASD, and an improvement in the response of the person with ASD (p. 1).

In the Social Story™'s original form, three types of short, direct sentences, *descriptive*, *directive*, and

perspective, were used. Descriptive sentences describe the social situation in terms of relevant social cues; directive sentences specify an appropriate behavioral response; and perspective sentences describe the feelings and responses of the student or others in the targeted situation (Gray, 2000b). Gray (2003) has suggested that perspective sentences should only be rarely used to describe the internal states of persons with autism.

Originally the use of illustrations was not generally recommended as "...they may be distracting, or a student may make an inaccurate interpretation of the situation based on the illustration" (Gray & Garand, 1993, p. 4). This recommendation was subsequently revised and the use of illustrations "...that reflect consideration of the age and personal learning characteristics of the person with ASD" (Gray, 2003, p. 5) is now suggested as beneficial to social understanding. The use of pictures may be viewed as consistent with other research on visual supports for children with autism (e.g., Dettmer, Simpson, Myles, & Ganz, 2000).

¹ Macquarie University Special Education Centre, Macquarie University, North Ryde, NSW, Australia.

² Correspondence should be addressed to: Georgina Reynhout, Macquarie University Special Education Centre, Macquarie University, 2109, North Ryde, NSW, Australia.; e-mail: georgina.reynhout@speced.sed.mq.edu.au

Over time, the suggested format for a Social Story™ has become more sophisticated. Gray (2003) now describes four basic sentence types: *descriptive*, *directive*, *perspective* and *affirmative*. “Affirmative sentences enhance the meaning of surrounding statements, often expressing a commonly shared value or opinion within a given culture” (Gray, 2003, p. 3). These four basic sentence types and a ratio that defines their frequency are the most important components of the Social Story™. A basic Social Story™ has a ratio of two to five descriptive, perspective and/or affirmative sentences for every directive sentence.

Two other types of sentences can be included: *control* and *cooperative* sentences. “Control sentences are statements written by a person with ASD to identify personal strategies to use to recall and apply personal information” and cooperative sentences are “...used to identify what others will do to assist the student” (Gray, 2003, p. 3). A ratio of zero to one directive or control sentences to between two and five descriptive and/or perspective sentences forms the complete Social Story™ ratio. Within a Social Story™, descriptive, perspective, control and directive sentences can be partial or complete. “Partial sentences encourage the student with ASD to make guesses regarding the next step in a situation, the response of another individual, or his own response” (Gray, 2003, p. 3).

There are several considerations to be made when writing a Social Story™ according to Gray’s format (2000a, 2000b, 2003). The author must adopt and maintain the perspective of the child for whom the story is written. The story should be well within the student’s comprehension level and clearly presented using vocabulary and print size that are appropriate for the student. Behavioral responses should be stated in positive terms, for example using “I will use my quiet voice” instead of “I will not shout.” It is suggested that Social Stories™ be used with students “...functioning intellectually in the trainable mentally impaired range or higher who possess basic language skills” (Gray & Garand, 1993, pp. 2–3).

Gray and Garand (1993) describe how Social Stories™ can be modified to suit a wide variety of purposes. For example, checklist or sequence Social Stories™ can teach and provide practice in the following of routines, and curriculum stories can be created by inserting pages involving academic activities, such as math, into the Social Story™. Skills related to specific curriculum objectives can then be practiced in the context of real life situations relevant to the student. Generic Social Stories™ can be used to

describe social situations frequently experienced by children with autism which can be individualized and adapted as the need arises (Gray & Garand, 1993).

A Social Story™ intervention can be implemented in one of three ways and the approach is dependent upon the particular needs and abilities of the child (Gray & Garand, 1993). If the child is able to read, the teacher can initially read the Social Story™ with the child, and the child thereafter reads the story independently. If the child is unable to read, the story can be recorded on cassette and the child taught to use the tape recorder, turning the page when prompted by an auditory signal. In this way the non-reader can “read” the story independently. The third method involves video modeling. The story is recorded onto videotape and sequences depicting the social situation and desired behavioral responses are matched to the story. Peers can read the Social Story™ to the child and when targeted situations arise, they are able to prompt the student to remember the Social Story™, thus reinforcing its use. Following the introduction of the story the student’s comprehension of the story is checked. This is a mandatory part of the process involving the teacher asking the student to respond verbally or in writing to verbal or written questions, or to fill in a checklist (Gray & Garand, 1993). Fading a Social Story™ can be achieved for example, by extending the time periods between readings, or by rewriting the Social Story™, omitting or revising selected sentences. Fading is highly individualized and again, dependent on the needs and abilities of the child (Gray & Garand, 1993).

Several possible explanations have been offered for the purported success of Social Story™ intervention for individuals with autism. One explanation involves the concept of “theory of mind” (ToM). Garfield, Peterson, & Perry (2001, pp. 495) define ToM as describing “...whatever knowledge guides propositional attitude attribution and the explanation and prediction of behaviour by means of inner states and processes.” Individuals with autism, lacking ToM, are unable to appreciate other people’s intentions, beliefs, needs and desires (Greenway, 2000). Research shows both linguistic (Happé, 1995) and social skills (Ozonoff & Miller, 1995) support the development of ToM. Individuals with autism, demonstrating significant impairments in both linguistic and social skills domains, consequently show ToM deficits (Garfield *et al.*, 2001). The perspective sentences in Social Stories™ are seen to address this social-cognitive deficit.

Another explanation involves the use of shared schemata or background knowledge. Rowe (1999) suggests that a Social Story™ builds a scaffold of understanding for a schema (mental representation) that an individual does not yet possess. Myles and Simpson (2001) describe Social Stories™ as providing access to a social skill's "hidden curriculum." This curriculum is based on the dos and don'ts innately understood and adhered to by everyone, except individuals with Asperger's syndrome.

Social Stories™ may also be seen to include aspects of purported good practice in ASD, being visual, permanent, written in simple language, based on careful assessment of the child, focused on key areas (for example, social interaction), and factual, focusing on the perceptions of others (Smith, 2001). In a survey of school staff, parents and caretakers trained in the Social Story™ approach, the majority reported it to be enjoyable, practical and effective (Smith, 2001).

Kuoch and Mirenda (2003) offered a limited review of research of 10 Social Story™ interventions studies prior to 2002 (including some case studies). Several limitations in the research were noted including non-conformity of many of the stories with recommended guidelines, flawed and weak research designs and confounding effects as additional interventions were often employed with Social Stories™. However, Kuoch and Mirenda (2003) included some studies that did not present learner outcome data, did not attempt to quantify outcome strength, examine generalization or maintenance and did not attempt to analyze whether intervention features were associated with differential outcomes. Thus, at this point, we do not have a comprehensive and detailed review of empirical studies on Social Story™ interventions.

Social Stories™ have inherent attractions in being relatively undemanding to implement and are reported to be applicable to a wide variety of behavior. The use of Social Stories™ has been popularized, widely discussed and recommended in the literature (e.g., Backman & Pilebro, 1999; Chapman & Trowbridge, 2000; Rowe, 1999; Simpson & Myles, 1998). Unfortunately, clinical popularity of interventions has not always been a good indicator of efficacy (see Dawson & Watling, 2000; Elder, 2002; Kasari, 2002; Kerrin, Murdock, Sharpston, & Jones, 1998; McWilliam, 1999). While there has been an increasing quantity of research examining Social Stories™ in recent years, no systematic and comprehensive review of empirical literature appears to have been conducted to date. The present paper addresses issues related to the nature and quality of studies

conducted, the participants, characteristics of the Social Story™ interventions with relation to suggested guidelines (see Gray, 2003; Gray & Garand, 1993), behaviors that have been targeted for intervention, the short-term efficacy of the technique, and generalization and maintenance. The paper includes a descriptive review of extant research as well as applying appropriate statistical analyses to relevant studies, including calculation of conventional effect sizes for group designs and single-subject meta-analysis.

IDENTIFICATION OF STUDIES

The following on-line databases were searched for sources appearing before December 2003: ABI/INFORM Global; Academic Research Library; Current Contents Connect; ERIC; Expanded Academic ASAP; First Search; Ingenta; Inspec; Kluwer Online; Proquest Education Complete; PsyARTICLES; psycINFO; Science Direct and ISI Web of Science. A manual search of all issues of *Journal of Autism and Developmental Disorders* was made for articles after 1990 and the reference sections of all located sources were reviewed for additional sources that did not appear in the on-line searches.

Only studies related to Social Stories™ were included in the review. Studies involving the use of social scripts, mutual storytelling and narrative therapy were excluded. Descriptive cases not including data (Del Valle, McEachern, & Chambers, 2001; Rowe, 1999), and studies not based on learner performance but teacher perception were excluded (Smith, 2001). Noting that the body of research remains relatively small, peer-reviewed journal articles and unpublished dissertations were included in the review. Eleven peer-reviewed journal articles (Bledsoe, Myles, & Simpson, 2003; Brownell, 2002; Hagiwara & Myles, 1999; Kuoch & Mirenda, 2003; Kuttler, Myles, & Carlson, 1998; Lorimer, Simpson, Myles, & Ganz, 2002; Norris & Dattilo, 1999; Rogers & Myles, 2001; Scattone, Wilczynski, Edwards, & Rabian, 2002; Swaggart *et al.*, 1995; Thiemann & Goldstein, 2001) and five dissertations (Cullain, 2002; Feinberg, 2002; Pettigrew, 1998; Romano, 2002; Staley, 2002) were identified.

PROCEDURES

A summary of each study was prepared addressing participants, research design, dependent variable

class, target behaviors, use of Social Story™ strategy, use of additional strategies, reliability, short-term findings, maintenance, and generalization, for each study that provided a verbatim copy of Social Stories™ used in intervention (totaling 31 stories from 13 of the 16 studies). Each sentence was independently rated for sentence type (i.e., descriptive, perspective, affirmative, directive, control and cooperative) using the guidelines provided by Gray (2003). Sentences were coded in multiple categories where this was considered appropriate. In the case of perspective sentences taking the format of “I feel happy when,” information provided in the second part of the sentence was considered to be supporting the first and was not coded elsewhere in terms of Gray’s (2000b, 2003) classification. Where perspective sentences were employed, they were coded as to whether they took the perspective of the target individual and/or that of others. One sentence type not described by Gray (2000b, 2003) was also coded. Sentences were also coded as to whether they specified a *consequence* of the actions of the target individual or others (e.g., “When I share my toys, my friends will want to play with me again”).

Both authors independently rated all stories to allow estimation of interrater reliability of sentence coding. An agreement was recorded when both raters indicated the presence of a particular coding category for a sentence and a disagreement was recorded when only one rater coded the category. Reliability was calculated by dividing agreements by the total of agreements and disagreements and multiplying by 100. Interrater reliability was 86.7% for descriptive sentences, 88.2% for perspective sentences, 85.7% for affirmative sentences, 93.5% for directive sentences, 85.7% for control sentences, 85.7% for cooperative sentences and 82.5% for consequence sentences. Where raters agreed on the presence of perspective sentence, agreements on whose perspective was taken was 96.8%.

Data on types of sentences were summarized in terms of the percentage of stories containing each sentence type and the mean percentage of each sentence type per story. In addition, the percentage of perspective sentences taking the perspective of the target individual, other individuals and both were calculated.

For group designs, presentation of effect sizes along with inferential statistics is recommended (American Psychological Association, 2001). Consistent with this recommendation, effect sizes were calculated on relevant studies by subtracting the

post-test mean of the control group from the mean of the experimental group and dividing by the standard deviation of the control group (McCartney & Rosenthal, 2000). Effect sizes were expressed positively when change occurred in the predicted direction and negatively when changes were opposite to those predicted. When targeting an increase in behavior, an effect size of .20 is small, .50 moderate, and .80 large; the usually accepted minimum clinically acceptable effect size for educational interventions is 0.33 (McCartney & Rosenthal, 2000).

Single-subject studies are typically interpreted by visual inspection of graphed data. Such visual inspection of single subject studies can be subjective and the objective aggregation of results is problematic. Calculation of the percentage of non-overlapping data (PND) has been suggested as an alternative (Scruggs, Mastropieri, & Casto, 1987). Using a graphical data plot of results from baseline to treatment, the percentage of non-overlapping data (PND) is the number of treatment data points that exceed the highest (or lowest, if appropriate) baseline data point, divided by the total number of treatment data points and multiplied by 100 (Scruggs *et al.*, 1987). While the procedure has been subject to criticism on several grounds (e.g., Salzberg, Strain, & Baer, 1987; Strain, Kohler, & Gresham, 1998; White, 1987), it has gained currency and has been applied across a range of areas (e.g., Didden, Duker, & Korzilius, 1997; Mathur, Kavale, Quinn, Forness, & Rutherford Jr., 1998; Scruggs, Mastropieri, Forness, & Kavale, 1988; Xin & Jitendra, 1999).

In the present analysis, a PND statistic was calculated for each study using the pooled number of non-overlapping data points across all subjects and all conditions. In addition, a PND statistic was calculated for each relevant story. A PND between 91 and 100 indicates a highly effective intervention, between 71 and 90 moderately effective, between 51 and 70 mildly effective, and between 0 and 50, non-effective (Mastropieri, Scruggs, Bakken, & Whedon, 1996).

RESULTS

A summary of the participants, research design and duration of study, target behaviors and settings, use of Social Story™ strategy and use of additional strategies was constructed and is presented in Table I. Table II provides descriptive information on reliability, short-term findings, study effect size or PND, maintenance, and generalization results.

Table 1. Subjects, Design and Duration, Target Behaviors, and Strategies

Study	Participants	Research Design		Target Behaviors		Use of Social Story™ strategy		
		(A)	(B)	(A)	(B)	(A) Construction	(B) Implementation	Additional Strategies
Bledsoe <i>et al.</i> (2003)	One boy, aged 13 with Asperger's syndrome and ADHD. Full scale IQ of 82.	(A) ABAB	(B) 7 days baseline, 5 days treatment; 5 day return to baseline, 4 day return to treatment.	(A) Appropriate social behavior (spills of food or drink, and napkin use).	(B) School dining area at lunchtime.	(A) One Social Story™ adheres to Gray's basic Social Story™ format.	(B) Social Story™ read by researcher. Social Story™ accessible to student following intervention. No comprehension session. Use of photos.	No additional strategies.
Brownell (2002)	Four boys aged 6–9 years, with reported autism. All boys were verbal and possessed at least pre-reading skills.	(A) ABAC (interventions counterbalanced across participants).	(B) Five days baseline, 5 days Social Story™ read; 5 days return to baseline, 5 days Social Story™ sung.	(A) Delayed echolalia. Response when given a direction for the first time. Frequency of using a loud voice.	(B) Classroom.	(A) Four Social Stories™; three stories adhering to Gray's basic Social Story™ format, one is inappropriately modified.	(B) Social Story™ read or sung by researcher. No comprehension session. Social Story™ presented in traditional and musical format.	No additional strategies.
*Cullain (2002)	Four boys and 1 girl, aged 6–10 years, with autism (diagnosis based on DSMV-IV manual). All participants used expressive language communication. All participants had histories of excessive behaviors.	(A) AB (reported as ABA but no data were collected during intervention phase).	(B) Social Story™ read to child two times per day for 5 days.	(A) Level of anxiety. Excessive behaviors such as: clinging to one friend, ignoring requests, verbal complaint, refusal to play, throwing objects, touching peers.	(B) Three participants in school setting; two participants in home setting.	(A) Five Social Stories™; one story adheres to Gray's basic Social Story™ format, two stories are appropriately modified and one is inappropriately modified.	(B) Social Story™ read by researcher or parent. Copy of Social Story™ taken home for daily reading. No comprehension session.	No additional strategies.

Table I. Continued

Study	Participants	(A) Research Design (B) Duration of study	(A) Target behaviors (B) Setting	Use of Social Story™ strategy	
				(A) Construction	(B) Implementation
*Feinberg (2002)	Twenty five boys and 9 girls, aged 8–13 years, with autism (diagnosis based on DSMV-IV manual, A-DOS-G, ADI-R and GARS). All participants had at least phrase speech. Participants had a mean full scale IQ of 71.29, SD 23.9.	(A) Experimental pre-test-post-test control-group. (B) Regular story read five times a day for one day and Social Story™ read five times a day for one day for control group; Social Story™ read five times a day for 2 days and once per day at home for 7 days for experimental group.	(A) Social skills: greeting behaviors, requesting to play a game, asking another person what they want to play, accepting another's choice of game. (B) "Game room".	(A) One Social Story™ inappropriately modified. (B) Social Story™ read by researcher for one group, and researcher and parent for other group. Copy of Social Story™ taken home for daily reading. Comprehension session included as a component of the intervention.	Verbal prompting.
Hagiwara and Myles (1999)	Three boys aged 7–9 years, with autism (diagnosis based on scores of 108, 96 and 82 obtained with ABC). Scores on PEP-R indicated developmental ages of 36, 26 and 40 months. All participants showed mild to moderate social skill problems and related behavior problems, and basic listening or written language skills.	(A) Multiple baseline across settings. (B) For participant I: 5 days at baseline, 19 days treatment (before morning snack); 13 days at baseline, 11 days treatment (before lunch); 19 days at baseline, 5 days treatment (after recess). For participant II: 4 days at baseline, 12 days treatment (before resource room); 8 days at baseline, 8 days treatment (before lunch); 12 days at baseline, 4 days treatment (after recess). For participant III: 3 days at baseline, 11 days treatment (lunch); 7 days at baseline, 8 days treatment (resource room); 7 days at baseline, 4 days treatment (classroom).	(A) Participants I and II: hand-washing skills. Participant III: on-task behavior. (B) For participant I, before morning snack, prior to lunch and after recess. For participant II before going to the resource room, before lunch, and after recess. For participant III lunch, resource room, and classroom.	(A) Only one Social Story™ of three provided; inappropriately modified. (B) Multi-media Social Story™ program operated by participants. No comprehension session. Social Story™ presented in computer based format. Use of visual symbols.	Verbal prompting. Physical assistance.

<p>Kuoch and Mirenda (2003)</p>	<p>Three boys, aged 3–6 years, with reported autism. Standard scores of 95, 44 and 107 obtained with PPVT-R.</p>	<p>(A) ABA for two participants. ACABA for one participant. (B) For one participant: 7 days at baseline, 5 days treatment, 5 days return to baseline (sharing). For one participant: 8 days at baseline, 8 days treatment, 13 days return to baseline (eating). For one participant: 8 days baseline, 7 days book and reminder, 7 days return to baseline, 8 days Social Story™, 8 days return to baseline (playing games).</p>	<p>(A) Sharing toys. Eating. Playing games. (B) One participant in home setting, one participant in preschool setting and one participant in school setting.</p>	<p>(A) Three Social Stories™ provided; one adheres to Gray's complete Social Story™ ratio, two appropriately modified basic ratio stories. (B) Social Story™ read by parent (who was a teacher) in the home setting, and by early interventionists in the preschool and school settings. No comprehension session. Use of visual symbols.</p>	<p>No additional strategies.</p>
<p>Kuttler <i>et al.</i> (1998)</p>	<p>One boy, aged 12, with reported autism, Fragile X syndrome and intermittent explosive disorder; with limited communication skills (vocalizations enhanced by communication book, manual signs and gestures). Scores on Callier-Azusa scale were cognition/60 months, receptive language/48 months, expressive language/24 months, and actual speech/18 months.</p>	<p>(A) ABAB (B) Five days at baseline, 5 days treatment, 3 days return to baseline and 6 days treatment.</p>	<p>(A) Precursors to tantrum behavior (inappropriate vocalizations and dropping to floor). (B) Self-contained classroom for children with autism.</p>	<p>(A) Two Social Stories™; one adheres to Gray's basic Social Story™ format, one inappropriately modified. (B) Each Social Story™ read by teacher. Social Story™ accessible to student following intervention. No comprehension session. Use of visual symbols.</p>	<p>Stickers/prize bag reinforcers incorporated into Social Story™.</p>
<p>Lorimer <i>et al.</i> (2002)</p>	<p>One boy, aged 5 years, with mild to moderate autism (diagnosis based on DSMV-IV manual), of average cognitive ability, with language skills commensurate with chronological age.</p>	<p>(A) ABAB (B) Seven days baseline, 7 days treatment.</p>	<p>(A) Precursors to tantrum behavior (interrupting vocalizations). Tantrum behavior. (B) Home setting.</p>	<p>(A) Two Social Stories™; both adhere to Gray's complete Social Story™ format. (B) Social Story™ read by parents and therapists. Social Story™ accessible to student following intervention. No comprehension session. Use of visual symbols.</p>	<p>No additional strategies.</p>

Table I. Continued

Study	Participants	(A) Research Design (B) Duration of study	(A) Target behaviors (B) Setting	Use of Social Story™ strategy	
				(A) Construction (B) Implementation	Additional Strategies
Norris and Dattilo (1999)	One girl, aged 8 years, with reported mild to moderate autism, able to verbally communicate, with cognitive functioning in the average, low average or mildly intellectually disabled range.	(A) AB (B) 5 days baseline; 15 days treatment.	(A) Appropriate social interactions. (B) Inappropriate social interactions. Absence of social interaction. (B) Lunch-time in school setting.	(A) No examples of Social Stories™ provided. (B) Social Story™ read by participant, with researcher. Social Story™ accessible to student following intervention. Comprehension session included as a component of the intervention. Randomized use of three different stories to provide interest. Use of visual symbols.	No additional strategies.
*Pettigrew (1998)	Forty-five boys and 24 girls, aged 3–4 years, with average intelligence (85+) and a language delay of at least 6 months as measured by norm-referenced tests.	(A) Experimental pre-test-post-test control and comparison-group design. (B) 20 sessions, 20 mins. each.	(A) Social skills: asking to play with a peer, sharing toys with a peer. (B) School setting.	(A) No examples of Social Stories™ provided. (B) No comprehension session. Stories personalized by inserting the first name of the child hearing the story. Use of visual symbols.	Scaffolding activities involving teacher modeling and rehearsal to practice social skills.
Rogers and Myles (2001)	One boy, aged 14 years, with Asperger Syndrome, learning difficulties and behavioral problems; able to verbally communicate.	(A) Pre-test/post-test (B) 5 days at baseline, 5 days treatment for Social Story™ one, Social Story™ two (comic strip) introduced day 12.	(A) Response to verbal directions. Tardiness to class. (B) School setting.	(A) Two Social Stories™; one adheres to Gray's basic Social Story™ format, one inappropriately modified. (B) Social Story™ read by participant, with researcher. No comprehension session. Use of visual symbols.	Treatment condition D, use of comic strip conversation.

*Romano (2002)	Five boys and 5 girls, aged 4–8 years, with reported autism, able to express needs/wants either verbally or non-verbally (using augmentative device, communication board, sign language or gestures).	(A) Experimental pre-test–post-test control-group design. (B) Five days at baseline, 30 days treatment.	(A) Inappropriate communication (echolalic, delayed or immediate), aggression (hitting, kicking and spitting) and inappropriate socialization (solitary play while in a group). (B) Classroom.	(A) No examples of Social Stories™ provided. (B) Social Story™ read by School Psychologist, teacher or child. No comprehension session, role-playing activity followed intervention. Each child received three Social Stories™ in random order on a daily basis. Use of visual symbols.	No additional strategies.
Scattone <i>et al.</i> (2002)	Three boys, aged 7–15 years, with reported autism and capable of communicating using speech.	(A) Multiple- baseline across participants. (B) Four days at baseline, 16 days treatment (tipping); 8 days at baseline, 12 days treatment (starling); 13 days at baseline, 7 days treatment (shouting).	(A) Disruptive behaviors (chair-tipping, shouting, inappropriate staring at girls). (B) School setting.	(A) Three Social Stories™; two adhere to Grays’ basic Social Story™ format, one adheres to Gray’s complete Social Story™ format. (B) Social Story™ read by teacher, teacher’s aide, or student. Social Story™ accessible to student following intervention. Comprehension session included as a component of the intervention.	Verbal prompting. An intervention to increase on-task behavior was implemented prior to, and during the study for one participant.
*Staley (2002)	Five males, aged 12–14 years, one with reported autism, two with Down Syndrome, one with Fragile X syndrome, and one with pervasive developmental disorder and seizure disorder. All participants had at least kindergarten level reading skills.	(A) Multiple baseline across participants. (B) 4–34 days baseline, 6 days Social Story™ (SS) intervention, 30 days SS with questions, 6 days reinforcer, 10 days return to SS with questions, 20 days return to reinforcer.	(A) Appropriate social behavior (chewing with mouth closed, and napkin use). (B) School cafeteria.	(A) Two Social Stories™; one adheres to Gray’s complete Social Story™ format; one is appropriately modified. (B) Social Story™ read by teacher, teacher’s aide, or doctoral student. Comprehension session included as a component of the intervention.	Verbal prompting. Behavior modeling. Edible reinforcers.

Table I. Continued

Study	Participants	(A) Research Design (B) Duration of study	(A) Target behaviors (B) Setting	Use of Social Story™ strategy	
				(A) Construction	(B) Implementation
Swaggart <i>et al.</i> (1995)	Three children, aged 7–11 years; one boy with pervasive developmental disorder, and one boy and one girl with reported autism. All participants had limited expressive language skills.	(A) AB B) 9 days baseline, 9 days treatment (greet); 51 days baseline, 10 + days treatment (aggression). 10 days baseline, 17/18 days treatment (sharing, grabbing, parallel play and aggression).	(A) Appropriate social interactions re: greeting, touching, aggression and ignoring. Appropriate social interactions re: sharing, grabbing, parallel play and aggression. (B) School setting.	(A) Four Social Stories™; two adhere to Gray's basic Social Story™ format, two inappropriately modified. (B) Social Story™ read by teacher or paraprofessional. No comprehension session. Use of visual symbols. Social Story™ paired with response-cost system for one participant.	Additional Strategies Verbal prompting and physical prompting, and Social Story™ paired with response-cost system used with one participant.
Thiemann and Goldstein (2001)	Five boys aged 6–12 years, with autism (diagnosis based on CARS scores), with impaired social communication, emerging/acquired word-identification skills and functional verbal communication, and 10 peers without disabilities.	(A) Multiple baseline across two to three social communication skills replicated across five triads. (B) Thirty minutes twice a week for triad; 10 min Social Story™ reading; 10 min. social interaction; 10 min. video feedback.	(A) Social skill in securing for attention, initiating comments, initiating requests and contingent responses. (B) School library.	(A) Four Social Stories™ used; one story only provided, adhering to Gray's basic Social Story™ format. (B) Social Story™ read by student with help from examiner or peers. No comprehension session. Use of visual symbols.	Verbal prompting. Direct social skills instruction involving written text cue rehearsal and role-play. Self-evaluation using video-feedback.

Note: ABC, Autism Behavior Checklist; A-DOS-G, Autism Diagnostic Observation Schedule-Generic; ADI-R, Autism Diagnostic Interview-Revised; CARS, Childhood Autism Rating Scale; DSM-IV, Diagnostic and statistical manual of mental disorders (4th edition); GARS, Gilliam Autism Rating Scale; PEP-R, Psycho-educational Profile-Revised; PPVT-R, Peabody Picture Vocabulary Test-Revised.
*Dissertation.

Table II. Reliability and Results

Study	Interobserver reliability	Procedural reliability	Reported short-term results	Effect size (ES) or PND	Maintenance and generalization
Bledsoe <i>et al.</i> (2003)	Calculated for 50% of sessions. Mean agreement 90%.	Procedural reliability not reported.	A reduction in food spillages. An increase in napkin use.	PND results from baseline to intervention for reduction in food spillages 11, for increase in napkin use 22. Study PND 16	Not reported
Brownell (2002)	Calculated for 40% of sessions. Mean reliabilities were 86%, 94%, 88% and 93%.	Procedural reliability not reported.	A reduction in target behaviors of each participant. Singing condition was significantly more effective than reading for one participant.	PND results from baseline 1(A) to story read (B) and baseline 2 (A) to story sung (C): Peter, 90; Brian, 88; Nathan, 80 and Justin, 90. Study PND 87.	Not reported.
*Cullain (2002)	Not reported.	Procedural reliability not reported.	A reduction in anxiety levels of three participants. A reduction in excessive behaviors for all participants.	PND results from baseline to post-intervention for three excessive behaviors: Douglas, 26; Sander, 80; Derek, 26; Kevin, 53 Beth 0. Study PND 37.	Maintenance not reported. Generalization: study conducted in two settings; three participants at school, two participants at home.
*Feinberg (2002)	Calculated for 25% of observations. Mean agreement 85%.	Procedural reliability not reported.	Increase in social skills scores for experimental group. No change in social skills scores for control group.	Mean effect size for intervention based on total social skills scores over six trials: 0.99 (highly effective).	Not reported.
Hagiwara and Myles (1999)	Calculated for 33% and 37% of sessions for two participants. 100% agreement. Calculated for 33% of sessions for one participant. Mean agreement 89%.	Procedural reliability not reported.	Improved hand-washing skills demonstrated. Participant I achieving 100% task completion in two settings; participant II achieving 92% task completion in one setting. Partially improved on-task behavior in two settings demonstrated by participant III.	PND results from baseline to intervention for hand-washing: Participant I, 39. Participant II, 37.5. PND results from baseline to intervention for on-task behavior: Participant III, 47. Study PND 40.	Maintenance not reported. Generalization: Study conducted in different school settings. One participant generalized skills to two new settings. One participant generalized skills to one new setting.
Kuoch and Mirenda (2003)	Calculated for each participant across a mean of 23.5% of all sessions. Mean agreement 97.9%.	Procedural reliability across all three participants was 98.4%.	A reduction in target behaviors for each participant.	PND results from baseline to intervention: Participant I, 0. Participant II, 100. Participant III, 75. Study PND 66.	Maintenance: decrease in targeted behaviors reported on return to baseline for two participants. Generalization: reported generalization of sharing and appropriate game playing skills for two participants.

Table II. Continued

Study	Interobserver reliability	Procedural reliability	Reported short-term results	Effect size (ES) or PND	Maintenance and generalization
Kuttler <i>et al.</i> (1998)	Calculated for 34% of observations. Mean agreement 93%.	Procedural reliability not reported.	A decrease in precursors to tantrum behavior during intervention. An increase in precursors to tantrum behavior post-intervention.	PND results from baseline 1 (A) to intervention (B) and baseline 2 (A) to intervention (B) during morning work time 100 and lunch time, 95. Study PND 95.	Maintenance not reported. Generalization: Study conducted in two school settings- morning work time and lunchtime.
Lorimer <i>et al.</i> (2002)	Calculated for 33% of observations. Mean agreement 96%.	Procedural reliability not reported.	A decrease in interrupting vocalizations and tantrum behavior during intervention. An increase in interrupting vocalizations and tantrum behavior post-intervention.	PND from baseline 1 (A) to intervention (B), and from baseline 2 (A) to intervention (B) results for frequency of interrupting vocalizations (B) 86; for frequency of tantrum behaviors 0. Study PND 43.	Not reported.
Norris and Dattilo (1999)	Calculated for 20% of baseline sessions, occurrence agreement 89–100%. Calculated for 25% of intervention sessions, occurrence agreement 88–100% (one at 64%).	Procedural reliability measures for each story were 100%.	No increase in appropriate social interactions. Overall decrease in inappropriate social interactions. Increase in no social interactions.	PND results from baseline 1 (A) to intervention (B) for frequency of appropriate social interactions 17, inappropriate social interactions, 8 and no social interactions, 25. Study PND 16.	Not reported.
*Pettigrew (1998)	Not reported.	Procedural reliability not reported.	An increase in pre-test-post-test SCBE scores for experimental and comparison groups.	Mean effect size for SCBE scores, experimental/control 0.34 (effective); experimental/comparison 0.26	Not reported.
Rogers & Myles (2001)	Not reported.	Procedural reliability not reported.	Improved response to verbal directions. Reduced tardiness to class.	Insufficient data provided.	Not reported.

*Romano (2002)	Calculated for 100% of sessions, occurrence agreement 90–100%.	Procedural reliability not reported.	Decrease in aggression, inappropriate communication and inappropriate socialization demonstrated by treatment and control groups from pre-test to post-test.	Mean effect size for aggression, inappropriate communication and inappropriate socialization was -1.59 , range -0.94 to -2.31 .	Maintenance: increase in inappropriate behaviors demonstrated from post-test to maintenance for treatment and control groups. Increase in inappropriate targeted behaviors demonstrated pretest to maintenance by control group. Decrease in inappropriate targeted behaviors demonstrated from pre-test to maintenance by treatment group. Generalization: not reported.
Scattone <i>et al.</i> (2002)	Calculated for 30% of observations. Mean agreement 93% for one participant. Occurrence agreement 100% for two participants.	Procedural reliability measures of 91% for one participant and 100% for 2 participants.	Decrease in disruptive behaviors (chair-ipping, shouting, inappropriate staring at girls).	PND results from baseline I (A) to intervention (B) for Kenny 100, John 92 and Howard 57. Study PND 88.	Not reported.
*Staley (2002)	Calculated for 15% of napkin use observations. Mean agreement 97%. Calculated for 18% of chewing observations. 100% agreement.	Procedural reliability not reported.	No increase in frequency in napkin use or chewing with mouth closed with Social Story™ used in isolation. Increase in frequency in napkin use and chewing with mouth closed when edible reinforcers used in isolation, or in conjunction, with Social Stories™.	Mean PND for participants I to V for napkin use, from baseline (B) to Social Story™ (S) 0, from (B) to Social Story™ and questions (SQ) 17 and from (B) to Social Story™, questions and reinforcer (R) 73. Results for participant I to V for chewing with mouth closed from (B) to (S)4, (B) to (SQ) 1, and (B) to (R)66. Study PND 30.	Not reported.
Swaggart <i>et al.</i> (1995)	Not reported.	Procedural reliability not reported.	Increase in appropriate social interaction re: greeting, touching, aggression and ignoring. Increase in appropriate social interactions re: sharing, grabbing, parallel play and aggression.	Mean PND for greeting behavior from baseline to intervention 88, for aggressive incidents from baseline to intervention 0. Study PND 24.	Maintenance not reported. Generalization: increased appropriate social interaction demonstrated across a variety of settings.

Table II. Continued

Study	Interobserver reliability	Procedural reliability	Reported short-term results	Effect size (ES) or PND	Maintenance and generalization
Thiemann and Goldstein (2001)	Calculated for 30% of sessions. Mean agreements for participants 92%, 89%, 92%, 93% and 90%.	Procedural reliability measured for 20% of sessions at 89%.	Increase in targeted social skills from baseline when treatment implemented.	Mean PND for Dan, 70, Greg 28, John 22, Casey 77, and Ivan 17. Study PND 45.	Maintenance: Limited maintenance in social skills demonstrated by three participants. Generalization: one participant demonstrated generalization of newly acquired social skills to one new setting.

Abbreviation: SCBE, Social Competence and Behavior Evaluation-Preschool Edition.
 *Dissertation.

A summary of the results of the analysis of the sentence types is presented in Table III.

Table IV provides a summary of aspects of story construction and implementation in relation to PND for individual stories for which that metric could be calculated. All stories involved one participant except those of Staley (2002), who used five participants for two Social Stories™, and Thiemann and Goldstein (2001), who used five participants for one story.

Social Story™ use comprises Social Story™ construction and implementation. Gray (2003) defines two Social Story™ types, those adhering to the basic Social Story™ ratio, and those adhering to the complete Social Story™ ratio. Gray (2003) has noted that directive sentences may not be necessary in some instances so stories with a lower than recommended ratio of directive to other sentences were considered to be appropriately modified. Based on analysis of the Social Stories™, four categories of Social Story™ were identified: (a) stories that adhered to the basic Social Story™ ratio; (b) stories that adhered to the complete Social Story™ ratio; (c) appropriately modified Social Stories™, where the number of descriptive, perspective, affirmative or cooperative sentences is greater than that recommended in the basic or complete ratios, but still acceptable according to Gray's recommendations (see Gray, 2003), and (d) inappropriately modified Social Stories™, where the number of directive or control sentences is greater than that recommended in the basic or complete ratios. The results of the analysis are presented in Table V.

Participants

Fifteen of the 16 studies involved a child or children with autism or Asperger's syndrome; of these children, 60 were boys and 17 were girls, aged 3–15 years. Pettigrew (1998) provided the only study

Table III. Social Story™ Sentence Analysis

Sentence type	Percentage of stories containing sentence type	Mean percentage of sentence type per story (range)
Descriptive	97	42 (0–91)
Perspective	87	23 (0–60)
Affirmative	22	3 (0–21)
Directive	100	28 (9–89)
Control	10	1 (0–14)
Cooperative	13	2 (0–29)
Consequence	90	24 (0–50)

Note: All percentages are rounded.

Table IV. Story Implementation and Construction Variables and PND

Social story™	Behavior	Percentage of sentence type:										SS ratio	COMP	Additional Strategies	PND
		DES	PER	AFF	DIR	CON	COOP	CONS							
Bledsoe <i>et al.</i> (2003)	Food spill, napkin use	25	23	0	25	0	0	50	Basic	No	None	16			
Brownell (2002; Brian)	Follow directions	50	17	0	33	0	33	Inappropriately modified	No	None	88				
Brownell (2002; Justin)	Using quiet voice	50	33	0	17	0	33	Basic	No	None	90				
Brownell (2002; Nathan)	Using quiet voice	57	14	0	29	0	29	Inappropriately modified	No	None	80				
Brownell (2002; Peter)	Echolalia	29	43	0	29	0	29	Inappropriately modified	No	None	90				
Cullain (2000; Sander)	Playing games	38	44	0	19	0	25	Basic	No	None	80				
Cullain (2002; Beth)	Talking with friends	13	60	0	27	0	13	Basic	No	None	0				
Cullain (2002; Derek)	Computer use	14	43	0	21	0	29	Complete	No	None	26				
Cullain (2002; Douglas)	Different friends	60	0	0	40	0	0	Basic	No	None	37				
Cullain (2002; Kevin)	Playing games	29	47	0	12	14	24	Complete	No	None	53				
Hagiwara and Myles (1999)	Hand washing	0	11	0	89	0	0	Inappropriately modified	No	Verbal prompts, physical prompts	38				
Kuoch and Miranda (2003; Neil)	Playing games	45	15	21	18	0	27	Basic (appropriately modified)	No	None	75				
Kuoch and Miranda (2003; Andrew)	Sharing	60	20	10	10	0	30	Basic (appropriately modified)	No	None	66				
Kuoch and Miranda (2003; Henry)	Eating	45	32	5	14	5	40	Complete	No	None	100				
Kuttler <i>et al.</i> (1998)	Behavior at lunch	38	13	0	50	0	25	Inappropriately modified	No	Reinforcers (stickers/prize bag)	90				
Kuttler <i>et al.</i> (1998)	Working for sticker	67	17	0	17	0	50	Basic	No	Reinforcers (stickers/prize bag)	100				
Lorimer <i>et al.</i> (2002)	Talking	53	13	0	20	13	7	Complete	No	None	43				
Lorimer <i>et al.</i> (2002)	Waiting	91	9	0	9	0	9	Basic (appropriately modified)	No	None	43				
Scattone <i>et al.</i> (2002; Howard)	Shouting	43	14	0	29	0	29	Complete	Yes	Verbal prompts	57				
Scattone <i>et al.</i> (2002; John)	Looking at girls	50	10	10	30	0	10	Basic	Yes	Verbal prompts	92				
Scattone <i>et al.</i> (2002; Kenny)	Sitting on chair	50	30	0	20	0	50	Basic	Yes	Verbal prompts	100				
Staley (2002)	Chewing	65	4	19	12	0	4	Complete	Yes	Verbal prompts, modeling, reinforcer (edibles)	32				
Staley (2002)	Napkin use	78	0	11	11	0	11	Basic (appropriately modified)	Yes	Verbal prompts, modeling, reinforcer (edibles)	39				
Swaggart <i>et al.</i> (1995)	Greeting	40	0	0	60	0	40	Inappropriately modified	No	Verbal prompts, physical prompts	88				
Swaggart <i>et al.</i> (1995)	Appropriate behavior	50	17	0	33	0	33	Basic	No	Verbal prompts, physical prompts, reinforcer (edibles)	0				

Note: DES, Descriptive; PER, Perspective; AFF, Affirmative; DIR, Directive; CON, Control; COOP, Cooperative; CONS, Consequence; SS ratio, Social Story™ ratio; COMP, Comprehension assessed. All percentages are rounded.

Table V. Summary of Story Analysis

Category of Social Story™	Number of stories
Basic	12
Complete	6
Appropriately modified	4
Inappropriately modified	9

not involving children with autism or Asperger's syndrome. The criteria for inclusion in this study was a language delay of at least 6 months and an age of 3–4 years (Pettigrew, 1998). In most instances, only a diagnostic label of autism or autism spectrum disorder was provided. Only three of the studies (Feinberg, 2002; Hagiwara & Myles, 1999; Thiemann & Goldstein, 2001) provided standardized data on the position of the participants on the autism spectrum. The communication skills of participants in the studies ranged from those who were non-verbal (using augmentative device, communication board, sign language or gestures), to those able to communicate verbally (Table I). In seven studies (Bledsoe *et al.*, 2003; Feinberg, 2002; Lorimer *et al.*, 2002; Norris & Dattilo, 1999; Pettigrew, 1998; Scattone *et al.*, 2002; Thiemann & Goldstein, 2001), reference was made to cognitive abilities of the participants; only four of these provided information based on the results of standardized tests (Bledsoe *et al.*, 2003; Feinberg, 2002; Scattone *et al.*, 2002; Thiemann & Goldstein, 2001). Only five studies (Bledsoe *et al.*, 2003; Brownell, 2002; Kuoch & Mirinda, 2003; Staley, 2002; Thiemann & Goldstein, 2001) provided description of the reading ability of the participants; four of these presented results of standardized tests (Bledsoe *et al.*, 2003; Kuoch & Mirinda, 2003; Staley, 2002; Thiemann & Goldstein, 2001). As the participant description provided in many studies consisted of little more than a diagnostic label, it was effectively impossible to determine if any specific participant characteristics were associated with intervention effectiveness.

Research Design

The experimental designs of the single-subject studies ranged in complexity from simple AB (Cullain, 2002; Norris & Dattilo, 1999; Swaggart *et al.*, 1995) to withdrawal (Bledsoe *et al.*, 2003; Brownell, 2002; Kuoch & Mirinda, 2003; Kuttler *et al.*, 1998; Lorimer *et al.*, 2002), to multiple baseline across participants (Scattone *et al.*, 2002; Staley, 2002), skills (Thiemann & Goldstein, 2001) or settings

(Hagiwara & Myles, 1999). Simple pre-test–post-test data was reported in some instances (Rogers & Myles, 2001; Swaggart *et al.*, 1995 for some participants). In a small number of studies group designs were employed including a pre-test–post-test control-group design (Feinberg, 2002; Romano, 2002) and a pre-test–post-test control and comparison-group design (Pettigrew, 1998). There was wide variation in study duration within and among the research designs employed in the studies.

Interobserver and Procedural Reliability

Twelve of the 16 studies (Table II) provided measures of interobserver reliability. In general, reliability measures in the studies ranged from 85–100%, calculated for between 15 and 100% of sessions (20% being the conventional minimum standard). On one occasion, in the study conducted by Norris and Dattilo (1999), the interobserver reliability dropped to 64%, resulting in a revision by the authors of their definitions of appropriate, inappropriate and absence of social interactions, to reflect a greater level of specificity. Procedural reliability measures were reported for only four studies (Kuoch & Mirinda, 2003; Norris & Dattilo, 1999; Scattone *et al.*, 2002; Thiemann & Goldstein, 2001).

Target Behaviors and Settings

The studies addressed various behaviors. Six studies targeted disruptive or challenging behaviors (Cullain, 2002; Kuttler *et al.*, 1998; Lorimer *et al.*, 2002; Romano, 2002; Scattone *et al.*, 2002; Swaggart *et al.*, 1995), nine studies targeted social skills (Bledsoe *et al.*, 2003; Feinberg, 2002; Kuoch & Mirinda, 2003; Norris & Dattilo, 1999; Pettigrew, 1998; Romano, 2002; Staley, 2002; Swaggart *et al.*, 1995; Thiemann & Goldstein, 2001), four studies targeted communicative behaviors (Cullain, 2002; Kuttler *et al.*, 1998; Lorimer *et al.*, 2002; Romano, 2002) and four studies targeted on-task behaviors (Brownell, 2002; Hagiwara & Myles, 1999; Kuoch & Mirinda, 2003; Rogers & Myles, 2001).

Twelve studies took place in a school setting (Bledsoe *et al.*, 2003; Brownell, 2002; Hagiwara & Myles, 1999; Kuttler *et al.*, 1998; Norris & Dattilo, 1999; Pettigrew, 1998; Rogers & Myles, 2001; Romano, 2002; Scattone *et al.*, 2002; Staley, 2002; Swaggart *et al.*, 1995; Thiemann & Goldstein, 2001), one study took place in a home setting (Lorimer *et al.*, 2002), two studies took place in a school and home setting (Cullain, 2002; Kuoch & Mirinda,

2003), and one study took place in a “game room” at an unidentified location (Feinberg, 2002).

Effectiveness

The authors of nine studies (Bledsoe *et al.*, 2003; Brownell, 2002; Cullain, 2002; Kuoch & Mirenda, 2003; Kuttler *et al.*, 1998; Lorimer *et al.*, 2002; Norris & Dattilo, 1999; Rogers & Myles, 2001; Romano, 2002; Scattone *et al.*, 2002) reported an appropriate reduction in target behaviors and eight studies (Bledsoe *et al.*, 2003; Feinberg, 2002; Hagiwara & Myles, 1999; Lorimer *et al.*, 2002; Pettigrew, 1998; Rogers & Myles, 2001; Swaggart *et al.*, 1995; Thiemann & Goldstein, 2001) claimed an appropriate increase in targeted behaviors from baseline to intervention, or from pre-test–post-test. Two studies showed no change in some behaviors when Social Stories™ were used in isolation (Norris & Dattilo, 1999; Staley, 2002) and two studies showed an actual increase in disruptive behaviors post-intervention (Kuttler *et al.*, 1998; Lorimer *et al.*, 2002).

Effect sizes were calculated for three studies (Feinberg, 2002; Pettigrew, 1998; Romano, 2002). Based on total social skills scores for six trials (Feinberg, 2002) a mean effect size of .99 was calculated, suggesting Social Story™ intervention was highly effective in increasing social skills in this study. Romano (2002) failed to provide standard deviations for the experimental and control groups at post-test. The mean effect size at maintenance was 1.59 (range .94 to 2.31) indicating that Social Story™ intervention was highly effective in decreasing the targeted behaviors for this group. These data should be treated with extreme caution as the post-test results indicated a massive experimental effect in the control group, which was reversed at maintenance, raising serious questions about experimental control in the study. Romano (2002) suggested that this might have been a product of contamination of the intervention as control group students may have overheard the Social Stories™. Analysis of the Pettigrew (1998) study (see Table II) revealed low mean effect sizes for the experimental-control contrasts of .34 and for the experimental-comparison contrasts of 0.26. In addition, these data also need to be treated with caution, as there were clear discrepancies between tabular and graphic data. Staley (2002) has also pointed out other serious discrepancies in the data and analysis.

A total PND value was obtained for 12 studies (i.e., Bledsoe *et al.*, 2003; Brownell, 2002; Cullain,

2002; Hagiwara & Myles, 1999; Kuoch & Mirenda, 2003; Kuttler *et al.*, 1998; Lorimer *et al.*, 2002; Norris & Dattilo, 1999; Scattone *et al.*, 2002; Staley, 2002; Swaggart *et al.*, 1995; Thiemann & Goldstein, 2001). The calculated mean PND was 43 (range 16–95). In some cases the data overlap may not provide an accurate measure of treatment effectiveness, for example, when baseline data shows an inappropriate trend, or when “floor or ceiling” effects occur (Scruggs *et al.*, 1987). Possible floor and ceiling effects were certainly present in the data examined in the present review (i.e., Bledsoe *et al.*, 2003; Cullain, 2002; Kuoch & Mirenda, 2003; Lorimer *et al.*, 2002; Norris & Dattilo, 1999; Staley, 2002; Swaggart *et al.*, 1995).

To evaluate the possible effects of such effects a further analysis was conducted. Data were eliminated for behaviors, subjects or phases that were expected to decrease that contained a zero data point and, similarly, data were eliminated for behaviors that were expected to increase and a baseline ceiling was evident in one data point (i.e., 100% of opportunities). When these data were excluded and the PND recalculated, a total PND of 51 (range 20–95) was obtained. The mean PND values place the Social Story™ in the non-effective range or at very best, in the low end of the mildly effective range (see Mastropieri *et al.*, 1996). The most striking feature of the data, however, was the degree of inconsistency, including variation in responses to intervention across participants or behaviors (see Table II).

Story Construction

Examples of Social Stories™ used for intervention were not provided in three studies (i.e., Norris & Dattilo, 1999; Pettigrew, 1998; Romano, 2002). Examples of Social Stories™ were provided for 13 of the 16 studies (i.e., Bledsoe *et al.*, 2003; Brownell, 2002; Cullain, 2002; Feinberg, 2002; Hagiwara & Myles, 1999; Kuoch & Mirenda, 2003; Kuttler *et al.*, 1998; Lorimer *et al.*, 2002; Rogers & Myles, 2001; Scattone *et al.*, 2002; Staley, 2002; Swaggart *et al.*, 1995; Thiemann & Goldstein, 2001). Detailed examination of these stories (see Table III) revealed that affirmative, control, and cooperative sentences were used infrequently while descriptive, perspective and directive sentences were the most commonly employed. In contrast, with Gray’s (2003) recommendation that perspective sentences should only occasionally be written from the viewpoint of the person with autism, 47% of perspective sentences

were written from this viewpoint, 47% were written from the viewpoint of others and 6% were written from both perspectives. Consequence sentences were not described by Gray (2003) but were identified in 90% of stories and contributed a mean of 24% (range 0–50%) of sentences per story.

It was of interest to examine whether specific aspects of story construction were associated with PND. Before proceeding with presentation of these data, it is appropriate to note that the analysis should be treated with caution as is based on a relatively small number of stories (26) on which PND can be calculated. Gray (2003) offers clear guidelines on how Social Stories™ should be constructed. Inspection of Table V indicates that analysis of the Social Stories™ reveals 10 (39%) of the stories deviate from recommended basic or complete Social Story™ ratios. One obvious question was whether deviation from Gray's guidelines (Gray, 2003) would differentially affect outcomes. From the data in Table IV, which is limited to stories where PND was available, it can be determined that the 14 stories with basic or appropriately modified basic ratios yielded a mean PND of 55.6. The six stories with complete story ratios yielded a mean PND of 51.8 and the six inappropriately modified stories, with a combination of less descriptive and more directive sentences, yielded a mean PND of 79.0.

In order to determine whether particular sentence types corresponded with differential outcomes, initially data in Table IV were ranked according to the frequency of each type of main sentence. Stories were then divided into approximate thirds and a PND statistic was calculated for the nine stories with the highest frequency of each sentence type, the eight stories that were middle ranked and the remaining nine sentences with the lowest rank. There was no evidence of an effect for perspective sentences with a PND of 64 for the highest ranked stories, 60 for middle ranked stories and 56 for the lowest ranked stories. The pattern was not consistent for directive sentences with a PND of 67 for the highest ranked stories, 45 for middle ranked stories and 66 for the lowest ranked stories. There was also inconsistency in the pattern for descriptive sentences with a PND of 59 for the highest ranked stories, 75 for middle ranked stories and 48 for the lowest ranked stories. There was some evidence of a pattern in PND for consequence sentences with a PND of 70 for the highest ranked stories, 65 for middle ranked stories and 47 for the lowest ranked stories. It was also interesting to note that the actual percentage of

consequence sentences was very low for the bottom third of stories (11.0% vs. 26.2% for middle ranked stories and 41.0% of highest ranked stories).

Story Delivery

In 15 of the studies, implementation of the Social Story™ was traditional, the Social Story™ being read to or by the participant(s) (Table I). In one study (Hagiwara & Myles, 1999) the Social Story™ was presented on a computer; in another study (Brownell, 2002) the Social Story™ was presented in both musical and traditional formats. In 11 studies the authors used visual symbols (Hagiwara & Myles, 1999; Kuoeh & Mirenda, 2003; Kuttler *et al.*, 1998; Lorimer *et al.*, 2002; Norris & Dattilo, 1999; Pettigrew, 1998; Rogers & Myles, 2001; Romano, 2002; Swaggart *et al.*, 1995; Thiemann & Goldstein, 2001) or photos (Bledsoe *et al.*, 2003) to enhance the information provided in the Social Story™ text.

There was variation in the number of Social Stories™ used in some studies. Norris and Dattilo (1999), for example, randomly used three different stories with one participant, to provide interest, targeting social interaction only. Romano (2002) used three different stories, each one targeting different behaviors (those associated with communication, aggression and socialization). Other variations in implementation procedures were also evident. The first variation involved the inclusion of some form of comprehension exercise, undertaken by the student following the reading of the Social Story™, as recommended by Gray and Garand (1993). In five stories from two studies (Scattone *et al.*, 2002; Staley, 2002) for which PND data were available, comprehension sessions were included as a component of Social Story™ intervention to enhance student understanding. From Table IV, it can be seen that the stories where authors reported a comprehension component yielded a mean PND of 64 and the 20 that did not yielded a PND of 60. The second relates to the student's access to the Social Story™ following its reading. In seven studies (Bledsoe *et al.*, 2003; Feinberg, 2002; Norris & Dattilo, 1999; Scattone *et al.*, 2002; Staley, 2002; Thiemann & Goldstein, 2001) the Social Story™ remained accessible to the student following its initial reading. During two studies the Social Story™ was taken home for the parents to read to the child (Cullain, 2002; Feinberg, 2002). Examination of Table IV reveals that the PND for the seven accessible stories was 64 and for the 19 remaining stories the PND was 63.

Use of Additional Strategies

Evaluation of the efficacy of Social Stories™ was confounded in some instances by the use of additional intervention strategies. Researchers used verbal and/or physical prompting in seven studies (Feinberg, 2002; Hagiwara & Myles, 1999; Rogers & Myles, 2001; Scattone *et al.*, 2002; Staley, 2002; Swaggart *et al.*, 1995; Thiemann & Goldstein, 2001) and tangible reinforcers such as stickers, or edibles were used in three studies (Kuttler *et al.*, 1998; Staley, 2002; Swaggart *et al.*, 1995). The study of Staley (2002) provided specific comparison of additional strategies and Social Stories™. Other strategies, including teacher modeling and rehearsal (Pettigrew, 1998; Thiemann & Goldstein, 2001) and self-evaluation using video-feedback (Thiemann & Goldstein, 2001), were also employed. From Table IV it can be determined that the 13 stories where additional strategies were employed yielded a PND of 59 and those that did not a PND of 62.

Maintenance and Generalization

Maintenance results were only reported for three of the 16 studies (Kuoch & Mirinda, 2003; Romano, 2002; Thiemann & Goldstein, 2001). Thiemann and Goldstein (2001) describes maintenance data across participants as “not compelling” and suggests a number of factors including length of training, skill difficulty, and adult prompt and visual cue fading to account for this finding. Romano (2002) collected post-intervention data on both the treatment and control group 6 weeks after discontinuation of the intervention for five consecutive days. Positive results were obtained for the treatment group, and negative results were obtained for the control group, that is there was an increase in inappropriate behaviors from pre-test and post-test to maintenance. As previously noted, however, the data from the Romano (2002) study needs to be interpreted circumspectly. Limited generalization in terms of targeted behaviors or newly acquired skills being demonstrated in more than one setting were reported in six studies (Cullain, 2002; Hagiwara & Myles, 1999; Kuoch & Mirinda, 2003; Kuttler *et al.*, 1998; Swaggart *et al.*, 1995; Thiemann & Goldstein, 2001). Cullain (2002) investigated Social Story™ intervention with three participants at school and two participants at home; the study by Kuttler *et al.* (1998) was conducted in two school settings, morning work time and lunchtime, and (Swaggart *et al.*, 1995) reported increased social interaction of three participants across a variety of settings.

Social Validity

Only three of the studies reviewed examined an aspect of social validity. Scattone *et al.* (2002) used the Intervention Rating Profile (Martens, Witt, Elliott, & Darveaux, 1985) with teachers to evaluate the acceptability of the treatment with the intervention falling well into the acceptable range. Thiemann and Goldstein (2001) asked 13 teachers and graduate students to view videotapes of social interaction recorded before and after intervention and to rate specific social behaviors for the target children and their peers; improvements in reciprocal social behaviors were reported by all raters. Hagiwara and Myles (1999) asked five educators and professors to verify the appropriateness of story construction.

DISCUSSION

Twelve of the 16 studies examined in this review used single subject designs. A major criticism of single subject designs is that they have low external validity. This weakness can be addressed by replication (Tawney & Gast, 1984). It is difficult to conduct replications and draw sensible conclusions about external validity if one is unsure of the characteristics of participants in existing research. A general criticism of the studies is that adequate descriptions of participant’s communicative and cognitive skills were not always provided. Gray and Garand (1993, p. 2) state “Social stories are most likely to benefit students functioning intellectually in the trainable mentally impaired range or higher who possess basic language skills.” Documentation of the level of cognitive and communicative functioning is therefore relevant and important, particularly considering the extreme variation possible in ASD. It is possible that the intervention is suited to participants with specific characteristics that will only be identified if the nature of the participant(s) is known. It is worthy of note that in all but one study (Pettigrew, 1998), participants were diagnosed with autism or Asperger’s syndrome. The efficacy of Social Stories™ in promoting behavior change in children with disabilities other than autism or Asperger’s is yet to be substantively researched. While almost all students in the studies reviewed were identified as being on the autism spectrum, quantification of the degree of autism was not provided in the vast majority of instances. Again, this may be an important variable impacting on the efficacy of Social Stories™.

Effect sizes in the small number of studies that used group designs were highly variable. The majority of studies used single subject designs and the PND statistic was calculated. The PND method sets a relatively high standard for evaluation of graphic data with the most extreme data point in baseline providing the point of comparison for intervention. Nevertheless, small *n* designs are not generally suited to detecting small treatment effects (unless data is exceedingly stable) and the practitioners are most interested in treatments with powerful clinical effects. The overall PND of 43 indicated Social Stories™ were an ineffective intervention according to the criteria of Mastropieri *et al.* (1996) and was only 51 (marginally effective) if zero or ceiling comparative data in baseline were removed. While some interventions evidence both high PNDs and reasonably clear demonstration of experimental control (e.g., Brownell, 2002; Kuttler *et al.*, 1998), effects were relatively modest (and/or inconsistent) in relation to baseline variation in other studies (e.g., Hagiwara & Myles, 1999; Norris & Dattilo, 1999), accounting for the relatively low PND figures. High levels of baseline instability in relation to the magnitude of intervention effects were certainly evident in a number of studies. In addition, Kuoch and Mirenda (2003) have noted that behavior change following Social Story™ intervention may not be fully reversible. Thus, it is possible that some studies using withdrawal components may not have been ideally suited to examining the efficacy of the intervention. There was also enormous variation in responses of specific participants to intervention. Considering the relatively small data set variation in PND across and within studies, the mean PND value should be treated with a high degree of caution. Some studies or interventions with specific students appeared to be highly effective with clear demonstrations of control and other studies were relatively ineffective, indicating the possibility that specific participant or intervention features influencing effectiveness.

Social Stories™ are claimed to be capable of addressing a variety of applications (Gray, 2003). The Social Stories™ used in the studies examined in the present review certainly addressed a variety of behaviors, supporting this claim. None of the Social Stories™ employed in the studies were of the curriculum type (see Gray & Garand, 1993). Thus, the efficacy of teaching academic skills through Social Story™ intervention remains unaddressed.

Considerable variation was seen in the use of Social Stories™ in the studies in terms of construction, implementation and use of additional strategies and these issues will be addressed in turn. Criticism of previous research has been offered by Kuoch and Mirenda (2003) on the basis that many stories failed to conform to the recommended construction. In a summary of the composition of Social Stories™ used in a selection of 10 case studies or experimental designs published in journal articles prior to 2003, Kuoch and Mirenda (2003) found five studies to contain Social Stories™ not conforming to applicable Social Story™ guidelines, although the impact of this on Social Story™ efficacy as determined by, for example, is not determined by the authors. Analysis of the Social Stories™ from the 16 experimental studies in the present review that included studies conducted prior to 2004, confirmed that a number of these also deviated considerably from the construction prescribed by Gray (2003). There was, however, no evidence that this impacted negatively on PND. In fact, the small number of stories that varied from Gray's (2003) guidelines (higher ratio of directive to descriptive sentences) seemed to be associated with considerably higher mean PND. The relationship between frequency of particular sentence types and PND was also of interest. There was no clear pattern of evidence that a higher absolute percentage of descriptive sentences impacted positively on PND. Stories in the middle third for descriptive sentence frequency were associated with considerably higher PNDs than those in the upper and lower third. There was preliminary evidence that the level of consequence sentences might affect efficacy, with stories demonstrating very low levels of consequence sentences being associated with lower PNDs. In hindsight, this is not surprising as strategies involving the systematic use of reinforcement has been identified as one of the most powerful interventions in special education (see Fuchs & Fuchs, 1986; Kavale & Forness, 1999; Skiba & Casey, 1985; Walberg & Wang, 1987) and have been consistently and demonstratively effective in facilitating behavior change with individuals with autism (see Heflin & Alberto, 2001; McConnell, 2002). The potential value of reinforcement is recognized by Gray (2003) who states:

Social stories have another purpose that is equally important: acknowledging achievement. In fact a child's first social story should describe a skill or

situation that is *typically successful and problem-free*. Written praise may be far more meaningful for children with ASD than its verbal counterpart. At least half the social stories developed for the child with ASD should bring attention to positive achievements (pp. 1–2).

In fact, consequences of actions were described in 90% of the Social Stories™ examined in the present study, suggesting the role of consequences may well be more important than suggested by Gray (2003). The exact role of consequences sentences in Social Stories™ awaits further empirical investigation but it is at least plausible that their role in identifying natural or artificial consequences of actions may be important. Analysis of the effects of Social Story™ composition should be interpreted carefully due to the small number of stories available, the fact that these variables were not experimentally manipulated and the percentages of various sentence types varied considerably. Nevertheless, there is some suggestion that some sentence types may impact on story efficacy although not necessarily in the way that might be predicted by Gray (2003).

In relation to reinforcement, it should be noted that three studies employed tangible reinforcers such as stickers or edibles (Kuttler *et al.*, 1998; Staley, 2002; Swaggart *et al.*, 1995) and a further study used self-evaluation with video-feedback (Thiemann & Goldstein, 2001). One study (Staley, 2002) is of particular interest in this regard. Five participants were read two Social Stories™ intended to increase their chewing with their mouth closed and increase their napkin use. The researchers also gave participants edible reinforcers for chewing with their mouth closed and for appropriate napkin use during lunch. Using a reversal and multiple-baseline designs to compare the effects of Social Stories™ and reinforcers, results clearly showed no effect of Social Stories™, but an almost immediate effect of reinforcers (Staley, 2002).

A range of other additional interventions, most frequently physical and verbal prompting, were employed in addition to Social Stories™. In studies where prompting and/or reinforcement are used as additional strategies, it becomes difficult to ascertain which of the treatments, prompting, reinforcement or Social Story™ is the critical component of the intervention, or whether a combination of these has the greatest effect. While there was no clear evidence of a difference between stories using additional strategies and those that did not in the present analysis, many confounding variables were evident. There is a very

clear research need to examine the extent to which Social Stories™ contributed additional impact to these well-validated intervention strategies.

The use of more than one Social Story™ to target one social situation, or the use of several stories to target several behaviors, the frequency with which a specific Social Story™ is reviewed, the inclusion of comprehension activities and the access of the participant to the Social Story™ outside the intervention phase are all variables that could affect intervention efficacy. The significance of these issues in terms of the effectiveness of the treatment requires further investigation.

Researchers used visual symbols in 11 of the studies examined. Originally, Gray and Garland (1993) advised against the use of illustrations, describing them as distracting and liable to make the student misinterpret the situation. This recommendation has subsequently been revised and the use of illustrations "...that reflect consideration of the age and personal learning characteristics of the person with ASD" (Gray, 2003, p. 5) is now suggested as beneficial to social understanding. The use of visual symbols is in keeping with current Social Story™ guidelines, and is consistent with other research on visual learning and children with autism (e.g., Dettmer *et al.*, 2000).

From the preceding discussion, it is obvious that Social Stories™ are multifaceted interventions. It is unclear from the present review that the prescribed (and complex) story construction is necessary to the efficacy of the intervention, which components are critical to effectiveness and whether Social Stories™ necessarily add to the effectiveness of other interventions. The confounding of Social Story™ interventions with other strategies is a problem in many existing studies (Kuo & Mirenda, 2003).

Maintenance and generalization were also inadequately addressed in the studies. These issues are of pivotal importance for children with ASD. On the basis of the limited information provided in the studies, it is not possible to draw conclusions about the efficacy of Social Story™ intervention in terms of maintenance and generalization. Programming for maintenance and generalization are typically essential components of an effective intervention. The inherent flexibility and portability of Social Stories™ may predispose them to be of use over time and in a number of settings.

Social validity usually refers to validation of goals, procedures and outcomes (Wolf, 1978) by

consumers, who can include participants, implementers, as well as members of the broader community. Hawkins (1991) and Schwartz and Baer (1991) have argued for more objective functional validation of intervention (habilitative validity), which may include comparison of performance with community standards. For example, Hughes, Harmer, Killian and Niarhos (1995) collected social comparison data on regular peers when teaching social initiation skills to students with intellectual disabilities. Thus, as well as demonstrating that the skills of interest could be taught, they also established that performance was within the broad functional range of regular peers. Smith (2001) has presented data suggesting that teachers viewed Social Stories™ as effective but objective data on learner performance was not provided, hence the study was not included in the present review. Only three of the studies reviewed addressed social validity, one examining acceptability of the intervention to teacher (Scattone *et al.*, 2002), the second the appropriateness of the Social Story™ construction and the third provided subjective validation of skill improvement in social skills by blind raters (Thiemann & Goldstein, 2001). Interestingly, none of the studies examined habilitative validity, although social comparative data would seem relevant to many of the skills addressed.

Some aspects of reliability measures in the studies were problematic. While interobserver reliability was reported in 12 studies but procedural reliability reported in only four studies. The relative simplicity of Social Stories™ intervention may lead to an assumption that procedural reliability does not need to be measured. Examination of Social Stories™ revealed deviation from recommended construction in a substantial proportion of instances and it would seem at least possible that there may have been deviations in the intended implementation. Reliability measures, including measures of procedural reliability, are a quintessential component of good research design; the lack of these undermines confidence in the research.

An important limitation of the existing review should be acknowledged. Results should be treated with caution given the modest data set available. In particular, analysis of specific study features should be treated circumspectly. This analysis was based on a subset of studies that yielded PND data and thus should be considered preliminary in nature. Interpretation is further limited by the degree of variation in story construction and implementation. Within the research reviewed there were several clear examples of

well-controlled studies suggesting the intervention was effective. Nevertheless, the intervention was not robustly effective across studies, participants and behaviors as reflected in mean PND. Consequently, the existing research does not demonstrate unequivocally that Social Story™ intervention is consistently effective in facilitating behavior change in children with autism. The majority of studies used single subject designs and the mean PNDs were marginal at best although there was substantial variation in results across studies, with some interventions proving very effective. This raises the possibility that undetermined factors (such as participant characteristics or story construction and implementation) may be important to the success of the intervention. This finding alone provides justification for further investigation of the strategy.

FUTURE RESEARCH

While there is a number of issues that need to be addressed in future research into Social Story™ intervention, several stand out as being of particular importance. Firstly a number of other procedural considerations should be addressed in future research on Social Stories™. Inadequate participation description in extant research has made it difficult to determine whether participant related variables moderate the effect of Social Story™ interventions. Adequate participant description, in terms of results of standardized tests, providing information pertaining to cognitive ability, position of the participant(s) on the autism spectrum, and language should be an integral component of future research investigating Social Story™ efficacy. The procedural integrity and social validity of Social Story™ interventions has not been examined in much of the extant research. Both these features are hallmarks of good research practice and should be routinely incorporated into future studies. In particular, habilitative social validity (Hawkins, 1991) should be examined to provide a meaningful framework for interpreting behavior change. In addition, there is some question whether the intervention is typically fully reversible and the use of reversal designs should be questioned. Repeated demonstrations of experimental control that are characteristics of strong small n research designs can be achieved using the multiple baseline design, without requiring interventions to be effectively reversible. Finally, very high levels of baseline instability in relation to intervention effects that was

evident in much of the research examined may be interpreted as indicating that future researchers might do well to concentrate more on establishing baseline stability.

Research on individuals with ASD and significant intellectual disabilities is extremely limited. Gray and Garand (1993) have suggested that the intervention may be suitable for individuals with a moderate level of intellectual disability (“trainable” in their terminology) who possess “basic language skills” but very few participants in the existing studies clearly fit this profile. Thus, the applicability of the intervention of students with ASD and significant intellectual disabilities needs to be explored further. Similarly, there is little research examining the potential for Social Story™ interventions to be used with student who do not present with ASD but do have major social skills deficits. Many interventions that have been successful with individuals who present with ASD (e.g., reinforcement, response prompting, visual supports) have also been effective with other groups. The extent to which Social Stories™ are effective with other populations or differentially effective with ASD remains to be explored.

There remains a question as to whether the construction of Social Stories™ contributes to the high variability in efficacy reported in the literature. Further research should be conducted to determine whether recommended story construction affects outcome or whether particular sentence types play a critical role. Noting that the recommended construction of Social Stories™ is not based on theoretical or empirical rationale (Kuo & Miranda, 2003), this issue should be considered a priority for future research. In addition, the comprehension component of Social Story™ intervention may be important to efficacy. Further research is necessary to explore this possibility.

The findings of many of the existing studies have been confounded by the use of Social Stores™ in combination with other empirically verified procedures, such as prompting and operant reinforcement. Thus, there is a pressing need for further research into the effectiveness of Social Stories™ in isolation.

Great variability was evident in the research reviewed in terms of story construction and implementation, including significant deviations from recommended practice. If such variation was evident in (presumably) relatively controlled research environments, questions are raised as to how the intervention might be being applied in the field. Noting that it could be argued that practice often outstrips the

research base, it would be of some interest to investigate the extent of Social Story™ use in the field as well as how the strategy is being applied.

Finally, research addressing issues of maintenance and generalization is extremely limited. This is surprising given that Social Story™ interventions are relatively simple and would seem suited to use over extended time periods as well as across persons and places. Future studies should be extended to incorporate investigation of maintenance and generalization a matter of priority.

CONCLUSION

A small corpus of research literature pertaining to the use of Social Stories™ raises many questions regarding their effectiveness in facilitating behavior change in children with autism. The present review of the literature highlights the need to adopt a systematic, rigorous scientific approach to future research if these questions are to be answered, and the efficacy of Social Stories™ determined.

REFERENCES

- American Psychological Association. (2001). *Publication Manual of the American Psychological Association* (5th ed.). Washington, DC: American Psychological Association.
- Backman, B., & Pilebro, C. (1999). Augmentative communication in dental treatment of a nine-year-old boy with Asperger syndrome. *ASDC Journal Of Dentistry For Children*, 66, 419–420.
- 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.
- Chapman, L., & Trowbridge, M. (2000). Social stories for reducing fear in the outdoors. *Horizons*, 11(3), 39–40.
- Cullain, R. E. (2002). *The effects of social stories on anxiety levels and excessive behavioral expressions of elementary school-aged children with autism*. Unpublished doctoral dissertation, The Union Institute Graduate College, Cincinnati, Ohio.
- Dawson, G., & Watling, R. (2000). Interventions to facilitate auditory, visual, and motor integration in autism: A review of the evidence. *Journal of Autism & Developmental Disorders*, 30, 415–421.
- Del Valle, P. R., McEachern, A. G., & Chambers, H. D. (2001). Using social stories with autistic children. *Journal of Poetry Therapy*, 14, 187–197.
- Dettmer, S., Simpson, R. L., Myles, B. S., & Ganz, J. B. (2000). The use of visual supports to facilitate transitions of students with autism. *Focus on Autism and Other Developmental Disabilities*, 15, 163–169.
- Didden, R., Duker, P. C., & Korzilius, H. (1997). Meta-analytic study on treatment effectiveness for problem behaviors with individuals who have mental retardation. *American Journal on Mental Retardation*, 101, 387–399.

- Elder, J. (2002). Current treatments in autism: Examining the evidence and clinical implications. *Journal of Neuroscience Nursing, 34*(2), 67–73.
- Feinberg, M. J. (2002). *Using social stories to teach specific social skills to individuals diagnosed with autism*. Unpublished doctoral dissertation, California School Of Professional Psychology, San Diego, California.
- Fuchs, L. S., & Fuchs, D. (1986). Effects of systematic formative evaluation: A meta-analysis. *Exceptional Children, 53*, 199–208.
- Garfield, J., Peterson, C., & Perry, T. (2001). Social cognition, language acquisition and the development of the theory of mind. *Mind & Language, 16*(5), 494–541.
- Gray, C. (2000a). *The new Social Story™ book*. (Illustrated edition). Arlington, TX: Future Horizons, Inc.
- Gray, C. (2000b). *Writing social stories with Carol Gray*. [Video-tape and workbook]. Arlington, TX: Future Horizons, Inc.
- Gray, C. (2003). Social Stories. Retrieved April 13, 2003, from <http://www.thegraycenter.org>.
- Gray, C., & Garand, J. D. (1993). Social stories: Improving responses of students with autism with accurate social information. *Focus on Autistic Behavior, 8*(1), 1–10.
- Greenway, C. (2000). Autism and Asperger syndrome: Strategies to promote prosocial behaviours. *Educational Psychology in Practice, 16*, 469–486.
- 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.
- Happe, F. G. E. (1995). The role of age and verbal ability in the theory of mind task performance of subjects with autism. *Child Development, 66*, 843–855.
- Hawkins, R. P. (1991). Is social validity what we are interested in? Argument for a functional approach. *Journal of Applied Behavior Analysis, 24*, 205–213.
- Heflin, L. J., & Alberto, P. A. (2001). Establishing a behavioral context for learning for students with autism. *Focus on Autism and Other Developmental Disabilities, 16*, 93–101.
- Hughes, C., Harmer, M. L., Killian, D. J., & Niarhos, F. (1995). The effects of multiple-exemplar self-instructional training on high school students generalized conversational interactions. *Journal of Applied Behavior Analysis, 28*, 201–218.
- Kasari, C. (2002). Assessing change in early intervention programs for children with autism. *Journal of Autism and Developmental Disorders, 32*, 447–461.
- Kavale, K. A., & Forness, S. R. (1999). *Efficacy of special education and related services*. Washington, DC: AAMR.
- Kerrin, R., Murdock, J., Sharpton, W., & Jones, N. (1998). Who's doing the pointing? Investigating facilitated communication in a classroom setting with students with autism. *Focus on Autism and Other Developmental Disabilities, 13*, 73–79.
- Kuoch, H., & Mirenda, P. (2003). Social Story™ interventions for young children with autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities, 18*, 219–227.
- Kuttler, S., Myles, B. S., & Carlson, J. K. (1998). The use of social stories to reduce precursors to tantrum behavior in a student with autism. *Focus on Autism and Other Developmental Disabilities, 13*, 176–182.
- 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*(1), 53–60.
- Martens, B. K., Witt, J. C., Elliott, S. N., & Darveaux, D. (1985). Teacher judgements concerning the acceptability of school based interventions. *Professional Psychology: Research and Practice, 16*, 191–198.
- Mastropieri, M. A., Scruggs, T. E., Bakken, J. P., & Whedon, C. (1996). Reading comprehension: A synthesis of research in learning disabilities. *Advances in Learning and Behavioral Disabilities, 10B*, 201–227.
- Mathur, S. R., Kavale, K. A., Quinn, M. M., Forness, S. R., & Rutherford, Jr., R. B. (1998). Social skills interventions with students with emotional and behavioral problems: A quantitative synthesis of single-subject research. *Behavioral Disorders, 23*, 193–201.
- McCartney, K., & Rosenthal, R. (2000). Effect size, practical importance, and social policy for children. *Child Development, 71*, 173–180.
- McConnell, S. R. (2002). Interventions to facilitate social interaction for young children with autism: Review of available research and recommendations for educational intervention and future research. *Journal of Autism and Developmental Disorders, 32*, 351–372.
- McWilliam, R. (1999). Controversial practices: The need for a re-accluturation of early intervention fields. *Topics In Early Childhood Special Education, 19*, 177–188.
- Myles, B. S., & Simpson, R. L. (2001). Understanding the hidden curriculum: An essential social skill for children and youth with Asperger Syndrome. *Intervention in School and Clinic, 36*, 279–286.
- 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.
- Ozonoff, S., & Miller, J. (1995). Teaching theory of mind: A new approach to social skills training for individuals with autism. *Journal of Autism & Developmental Disorders, 25*, 415–435.
- Pettigrew, J. D. C. (1998). *Effects of the modeling of verbal and non-verbal procedures for interactions with peers through social stories and scaffolded activities on the social competence of 3- and 4-year-old children with specific language impairments*. Unpublished doctoral dissertation, Texas Woman's University, Denton, Texas.
- Rogers, M. F., & Myles, B. S. (2001). Using social stories and comic strip conversations to interpret social situations for an adolescent with Asperger Syndrome. *Intervention in School and Clinic, 36*, 310–313.
- Romano, J. (2002). *Are social stories effective in modifying behavior in children with autism?* Unpublished doctoral dissertation, Fairleigh Dickinson University, Teaneck, New Jersey.
- Rowe, C. (1999). Do social stories benefit children with autism in mainstream primary schools?. *Special education forward trends, 26*(1), 12–14.
- Salzberg, C. L., Strain, P. S., & Baer, D. M. (1987). Meta-analysis for single-subject research: When does it clarify, when does it obscure?. *Remedial and Special Education, 8*, 43–48.
- 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.
- Schwartz, I. S., & Baer, D. M. (1991). Social validity assessments: Is current practice the state of the art?. *Journal of Applied Behavior Analysis, 24*, 189–204.
- 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.
- Scruggs, T. E., Mastropieri, M. A., Forness, S. R., & Kavale, K. A. (1988). Early language intervention: A quantitative synthesis of single subject research. *Journal of Special Education, 22*, 259–283.
- Simpson, R. L., & Myles, B. S. (1998). Aggression among children and youth who have Asperger's syndrome: A different population requiring different strategies. *Preventing School Failure, 42*, 149–153.
- Skiba, H. M., & Casey, A. (1985). Intervention for behaviorally disordered students: A quantitative review and methodological critique. *Behavioral Disorders, 10*, 239–252.

- Smith, C. (2001). Using social stories to enhance behavior in children with autistic spectrum difficulties. *Educational Psychology in Practice, 17*, 337–345.
- Staley, M. J. (2002). *An investigation of social-story effectiveness using reversal and multiple-baseline designs*. Unpublished doctoral dissertation, University of Kansas, Lawrence, Kansas.
- Strain, P. S., Kohler, F. W., & Gresham, F. (1998). Problems in logic and interpretation with quantitative syntheses of single-case research: Mathur and colleagues (1998) as a case in point. *Behavioral Disorders, 24*, 74–85.
- Swaggart, B., Gagnon, E., Bock, S. J., Earles, T. L., Quinn, C., Myles, B. S. *et al.* (1995). Using social stories to teach social and behavioral skills to children with autism. *Focus on Autistic Behavior, 10*, 1–16.
- Tawney, J. W., & Gast, D. L. (1984). *Single subject research in special education*. New York: Merrill.
- Thiemann, K. S., & Goldstein, H. (2001). Social stories, written text cues, and video feedback: Effects on social communication of children with autism. *Journal of Applied Behavior Analysis, 34*, 425–446.
- Walberg, H. J., & Wang, M. C. (1987). Effective educational practices and provisions for individual differences. In M. C. Wang, M. C. Reynolds, & H. J. Walberg (Eds.), *Handbook of special education: Research and practice – Learner characteristics and adaptive education* (Vol. 1, pp. 113–128). Oxford: Pergamon.
- White, O. R. (1987). Some comments concerning “The quantitative synthesis of single-subject research”. *Remedial and Special Education, 8*, 34–39.
- Wolf, M. W. (1978). Social validity: The case for subjective measurement or how applied behavior analysis is finding its heart. *Journal of Applied Behavior Analysis, 11*, 203–214.
- Xin, Y. P., & Jitendra, A. K. (1999). The effects of instruction in solving mathematical word problems for students with learning problems: A meta-analysis. *Journal of Special Education, 32*, 207–225.

Copyright of *Journal of Autism & Developmental Disorders* is the property of Springer Science & Business Media B.V. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.