Examination of Social Story Format on Frequency of Undesired Behaviors

Robyn Thompson, Ph.D.
University of Utah

Susan S. Johnston, Ph.D.
University of Utah

Abstract

The purpose of this investigation was to explore whether a difference existed between the effectiveness of paper-based format and tablet computer-based format Social Story interventions on frequency of undesired behaviors. An adapted alternating treatment design was implemented with four children with autism spectrum disorder (ASD). Data regarding frequency of undesired behaviors during target activities as well as social validity data regarding the perceived effectiveness and efficiency of the interventions were collected. Results revealed that Social Stories presented in both paper-based and tablet computer-based formats were effective in decreasing undesired behaviors when compared to baseline conditions, and that a notable difference did not exist between the effectiveness of the paper-based and tablet computer-based formats. Social validity data provided strong support for the use of both paper-based and tablet computer-based Social Story interventions in early childhood classrooms. Behavioral outcomes did not vary across paper-based and tablet computer-based formats. However, social validity data revealed that interventionist preference did vary across formats suggesting that interventionist preference should be considered when developing and implementing Social Story interventions.

Key Words: Early Childhood Special Education, Social Stories

Examination of Social Story Format on Frequency of Undesired Behaviors

Educators and practitioners are in search of interventions that are effective and that will improve the quality of life for persons with autism spectrum disorder (ASD; Horner, Carr, Strain, Todd, & Reed, 2002). One education based intervention is Social Stories™. Introduced by Carol Gray in 1991, Social Stories are used to share, with a student, relevant information including where and when a situation takes place, who is involved, what is occurring, and why (Gray, 2004). Social Stories provide explanations and appropriate behavioral options for specific situations, and may help individuals with ASD compensate for deficits in social perceptions (Gray & Garand, 1993). An emerging body of evidence has been established that examines the efficacy of Social Stories in helping children with characteristics of ASD achieve functional goals and enhance abilities. Quantitative literature reviews reveal that the overall effectiveness of Social Stories varies widely and authors of these reviews suggest that plausible explanations for this wide range in effectiveness include variability in research methodology across studies as well as variability in the implementation of Social Stories interventions (see Kokina & Kern, 2010; Leaf et al, 2015; Reynhout & Carter, 2011; Test, Richter, Knight, & Spooner, 2011 for reviews).

One aspect of Social Story interventions that has varied across investigations is presentation format. Social Story interventions have been presented in a number of ways, including; paper-

based booklet formats, computer-based formats, song-based formats, and iPad-based formats (e.g., Brownell, 2002; Chan & O'Reilly, 2008; Chan et al, 2011; Vandermeer, Milfor, Beamish, & Lang, 2013). Recent studies are beginning to examine whether or not the presentation format influences outcomes. Brownell (2002) compared the effectiveness of paper-based Social Stories and song-based Social Stories with four individuals ages 4-6 with a diagnosis of ASD. Outcomes revealed that the intervention resulted in a reduction of undesired behaviors for both conditions. However, the frequency of undesired behaviors was lower and more stable under the song-based condition. Similarly, Mancil, Hayden and Whitby (2009) compared the effects of paper-based and computer-based formats on decreasing the negative behaviors of three elementary school-aged participants with ASD. Results demonstrated a decrease in the frequency of undesired behaviors in both conditions. However, the frequency of undesired behaviors was lower when the Social Stories were presented in the computer-based format (Mancil et al., 2009). Although preliminary, the outcomes of Brownell's (2002) and Mancil et al.'s (2009) studies suggest that Social Story outcomes may be enhanced when presented in alternative formats. However, additional studies are needed to investigate the impact of mode of presentation on the effectiveness of Social Story interventions. Specifically, given the increased availability and use of tablet computers in classroom settings, research should compare outcomes when using paper-based and tablet computer-based Social Stories.

The proposed study is designed to answer the following questions:

1. Is a Social Story presented in a paper-based format more effective in decreasing undesired behaviors related to a target situation than a no-intervention baseline condition?
2. Is a Social Story presented in a tablet computer-based format more effective in decreasing undesired behaviors related to a target situation than a no-intervention baseline condition?
3. Is there a difference between the efficiency and effectiveness of paper-based format and tablet computer-based format Social Story interventions?

Method

Participants

Four early childhood aged students, Adan, Brad, Daniel, and Ethan (pseudonyms), who (a) had a diagnosis of an ASD, demonstrated characteristics of an ASD as determined by school classroom placement personnel, or demonstrated characteristics of challenging behaviors that interfered with classroom learning and participation, (b) had normal (or corrected to normal) vison and hearing, (c) had an interest in books as defined by a score of four or above on all questions on a teacher completed Preschool Book Interest 6-point Likert Scale (Kuoch & Mirenda, 2003), (d) had no prior exposure to Social Stories in the classroom setting, and (e) engaged in at least two challenging behaviors that occurred during different classroom activities, and served different behavioral functions, participated in this investigation. Teachers in the identified early childhood programs nominated children who met the inclusion criteria. A letter was then sent to the parents of the nominated children to request permission for their child to participate in the study. After consent was obtained, the researchers conducted a file review to obtain demographic information and observations to complete a functional assessment of challenging behaviors. Table 1 provides

a summary of participant characteristics, each participant's average score from the Preschool Book Interest Scale, and a summary of functional assessment data collected for each participant.

Experimental Design

An adapted alternating treatment design (AATD) was used to (a) examine the effect of Social Stories as a tool to decrease undesired behaviors in early childhood special education settings, and (b) determine if the story format (paper-based book format or tablet computer-based format) influenced outcomes. When implementing an AATD, the researcher alternates the presentation of two interventions or treatment conditions with individual participants to address two different behaviors that are equally difficult for a target student to participate in or achieve, but are functionally independent of one another (McDonnell, Jameson, & Rose, 2011). For the purposes of this study, the independent variables were the paper and iPad-based Social Stories, and the dependent variable was the frequency of undesired behaviors during identified activities. Intermittent baseline probes were used to measure frequency of challenging behaviors prior to the introduction of intervention. Daily intervention probes were used to evaluate the impact of the intervention. Intermittent probes of post-intervention behaviors served as a maintenance check to determine whether experimental effects were durable over time.

Setting

This study was conducted in four different early childhood special education classrooms. Two participants (Adan and Brad) attended separate, self-contained preschool classrooms in the same private school for children with characteristics of ASD. A third participant (Daniel) attended a self-contained special education classroom in a different private school for children with characteristics of ASD. The fourth participant (Ethan) attended an inclusive preschool classroom in a public school that served children with and without disabilities.

Table 1

Participant Age, Diagnosis, Assessment Data, and Book Interest Average Score

Participant Name	**Age at Start of Study**	**School-based Assessment Data**	**Book Interest Score**	**Target Activity**	**Behavior Function**
Adan	4 years, 7 months (55 months)	Learning Accomplishment Profile (Revised Edition) Gross Motor: 38 (approx. 63 months), Fine Motor: 26 (approx. 51 months), Pre-writing: 22 (approx. 54 months), Cognitive: 14 (approx. 40 months), Language: 14 (approx. 43 months), Self Help: 35 (approx. 51 months), Personal Social: 19 (approx. 47 months)	5.25	Circle Time	Escape undesired tasks and/or activities
				Small Group	Obtain attention from staff and peers
Brad	3 years, 8 months (44 months)	Early Learning Accomplishment Profile Gross Motor: 85 (approx. 29 months), Fine Motor: 58 (approx. 21 months), Cognitive: 65 (approx. 16 months), Language: 25 (approx. 13 months), Self Help: 37 (approx. 25 months), Personal Social: 32 (approx. 25 months)	4.75	Small Group A	Escape undesired tasks and/or activities
				Small Group B	Obtain attention from staff and peers
Daniel	6 years, 10 months (82 months)	Comprehensive Assessment of Spoken Language- Percentile rank: Basic Concepts: 66%**,** Antonyms: 97%**,** Sentence Completion: 13%**,** Paragraph Comprehension: 50%**,** Pragmatic Judgment: 9%	5.75	Lunchtime	Obtain attention from staff
				Bathroom	Obtain access to desired items/activity
Ethan	4 years, 10 months (58 months)	Battelle Developmental Inventory (2nd ed.)- Percentile Rank: Adaptive: 61%, Personal-Social: 53%, Motor: 50%, Cognitive: 73%; Communication: not reported	5.75	Circle Time	Obtain attention from staff and peers
				Small Group	Obtain access to desired object or activity

Adan and Brad attended their preschool program for six hours a day, four days per week. Daniel attended his early childhood program for six hours a day, four days per week; and for three and a half hours, one day per week. Ethan attended his preschool program for two hours a day, three days per week. Five adults staffed Adan's classroom of eight children, five adults staffed Brad's classroom of nine children, five adults staffed Daniel's classroom of ten children, and four adults staffed Ethan's class of 11 children.

The teachers and paraeducators across the four classrooms had varying levels of education. The lead teacher in Adan's class had a master's degree in early childhood special education. The lead teacher in Brad's class had a bachelor's degree in psychology. The lead teacher in Daniel's class had a bachelor's degree in early childhood education. Finally, the lead teacher in Ethan's class had a bachelor's degree in early childhood special education. The paraeducators across each of the four classrooms had varying levels of education ranging from a high school diploma to a college bachelor's degree.

Classroom consultants also supported the staff/children in each of the participants' classrooms. A speech-language pathologist served as an itinerant classroom consultant for Adan and Brad. Speech-language pathologists and occupational therapists served as itinerant classroom consultants for Daniel and Ethan.

Interventionist

In order to control for the potential effect of differences across interventionists, the first author served as the interventionist for all participants as well as a researcher for this investigation. The interventionist was not employed by the early childhood classrooms where the study was conducted but had 8 years of experience as a school-based occupational therapist as well as a Master's degree in early childhood special education.

Materials

For each participant, two Social Stories were written. The two Social Stories addressed different undesired behaviors that served different functions and occurred in the context of different activities (see procedures) and were written using the guidelines "Social Stories 10.1" established by Gray (2010). The interventionist used data from functional behavioral assessment when writing the Social Stories. All Social Stories were examined by three professionals with advanced training in writing Social Stories (one early childhood special educator, one school-based occupational therapist, and one clinic based/school-based occupational therapist) in order to confirm that each Social Story met Gray's criteria (Gray, 2010) and provided feedback on the stories. The researchers revised the Social Stories based upon the feedback provided. After each Social Story was written, the stories were transferred into a book-like format using Keynote software (Apple, 2013). The researchers also created two video clips of typically developing early childhood aged children engaging in the desired behaviors discussed in each Social Story. These video clips were used in the iPad-based Social Story interventions. After developing, illustrating, and inserting the video clips for each Social Story using Keynote, the researchers randomly selected one story for each participant to be presented in a paper-based format and one story to be presented in an iPad-based format.

The story for each participant that was selected for the paper-based format was downloaded from the iPad on which it was created to a desktop computer, and then printed in color onto 8 ½" X 11" white paper. The researchers then inserted the printed pages into clear page protector covers and assembled the pages in a booklet format. The story that was selected for the iPad-based format included two video clips. These video clips were embedded into the relevant pages of the Social Story. An iPad tablet computer was used when presenting the iPad-based Social Story intervention to each participant.

The characteristics of the paper-based and iPad-based Social Stories that were the same were: (a) use of Gray's criteria, (b) Social Story criteria verified by three professionals, (c) font size/style/color, (d) number of sentences (i.e., plus/minus two sentences), (e) Flesh Reading Ease score between 80-100, (f) number of clip art illustrations, and (g) number of photo illustrations. The characteristics of the paper-based and iPad-based Social Stories that were different were: (a) the use of two video modeling clips in the iPad-based Social Story, and (b) turning story pages by swiping a finger across the iPad screen versus manually turning the paper-based story pages.

Procedures

Identification of target activities and related behaviors. For each participant, the researchers conducted classroom observations and discussions with classroom teachers to identify two different activities in which the participants engaged in different undesired behaviors that were thought to serve different functions. Once identified, the researchers conducted a functional behavior assessment for each activity/behavior (O'Neill et al., 1997) to identify the functions of the undesired behaviors as well as desired replacement behaviors. The undesired behaviors and desired behaviors were reviewed and evaluated for functional and procedural equivalence and the researchers compared the unrelated desired behaviors and unrelated undesired behaviors for each participant to ensure similar rate of reinforcement, quality of reinforcement, level of response effort, immediacy of reinforcement (Mace & Roberts, 1993), activity type (i.e., teacher directed, student directed, free play, other), and student engagement (i.e., active, passive, other) in order to increase believability that the treatment affected the outcomes, rather than differences between the difficulty of the desired and/or undesired behaviors (McDonnell et al., 2011). Once the researchers determined that a participant's desired and undesired behaviors were functionally independent and equally difficult, the researchers proceeded to the baseline phase of the study for that participant.

Baseline phase. Baseline data were collected on the occurrence of the specified undesired behaviors during each of the two activities. Following collection of baseline data, the researchers compared the mean rates of undesired behaviors across the two activities to ensure that the rates were comparable. The researchers then randomly assigned the paper-based Social Story to one activity and the iPad-based Social Story to the remaining activity and proceeded to the comparison phase.

Comparison phase. Each participant received daily counterbalanced presentations of the paper-based and iPad-based interventions during the comparison phase to control for potential ordering effects. The length of time between presentations of the two daily interventions was approximately 1 hour.

Reading of the Social Stories took place in a one on one interaction between the interventionist and participant in areas identified by the lead teacher as ones that would not be distracting for the participant or the other students in the classrooms. The intervention occurred immediately prior to the identified classroom activity. During both the paper-based and iPad-based interventions, the interventionist read the Social Story directly to the participant. The participant was given the opportunity to turn the pages of the story, touch the illustrations, and make comments about the story as the interventionist read the Social Story. In addition, during the iPad-based intervention, the participant was given the opportunity to view the embedded video clips. Following the reading of each Social Story, the interventionist reviewed the story with the participant to assess and increase the participants' comprehension of the concepts and circumstances presented in the story (Gray, 2010; Kokina & Kern, 2010). Upon completion of the review, the interventionist helped the participant transition directly to the identified activity.

The comparison phase continued until a participant demonstrated either a differential effect or comparable level of performance between the two interventions (McDonnell et al., 2011). Data from the comparison phase was used to determine the next phase of the investigation. If there was a comparable level of performance between the two conditions, then the participant entered a flipped intervention phase. If there was a differential effect between the two conditions, then the participant entered the most effective intervention phase (McDonnell et al., 2011).

Flipped or most effective intervention phase. If it was determined that there was a comparable level of performance between the two interventions, the interventions were flipped so that the activity that was paired with the paper-based Social Story in the comparison phase was now paired with the iPad-based Social Story intervention and the activity that was paired with the iPad-based Social Story was now paired with the paper-based Social Story.

If the comparison phase data revealed a difference between the rates of behavior when using paper-based and iPad-based Social Story interventions, then the "most effective" intervention was used across both activities. Specifically, if the paper-based format was the most effective intervention for a participant, then the pages from the Social Story that was originally delivered in the iPad-based format would be printed onto paper, inserted into clear page protector covers, and assembled into a booklet format. Conversely, if the iPad-based story was the most effective intervention for a participant, then the Social Story that was originally delivered through a paper-based format would be delivered via the iPad.

Maintenance. Following the completion of the flipped/most effective intervention phase, maintenance data were collected approximately one time per week for four weeks for each participant. During maintenance, all interventions were discontinued. Maintenance data were collected in the context of the same activities as baseline and intervention sessions.

Data Collection

Intervention strategy. Data to assess the effectiveness and efficiency of the intervention strategies were collected during baseline, comparison, flipped/most effective intervention, and maintenance phase sessions. The researchers observed each participant during the identified activities and collected data using a frequency counting method to gather information on the

occurrence of undesired behaviors. The frequency of undesired behaviors was converted to a rate of undesired behaviors per minute to allow for comparison across sessions in which the length of the activity varied slightly from day to day.

Social validity. Data to assess the acceptability, perceived effectiveness, and perceived efficiency of the intervention strategies was collected via Goal Attainment Scaling, video ratings, and teacher completed surveys.

Goal attainment scaling. Progress made toward each participant's desired behaviors was documented, quantified, and compared using Goal Attainment Scaling (GAS; Roach & Elliot, 2005). GAS uses interviews during goal-setting and post treatment sessions to help determine indicators of a person's progress that can be challenging to assess using available standardized measures. Each lead teacher used the GAS to rate their participant's progress toward the specified goals following the final week of the study. The lead teachers did not view data or graphs related to the intervention outcomes, but they did observe at least one session of the baseline, comparison, flipped intervention, and maintenance phases of the study before providing ratings on the GAS.

Video ratings. Video of the participants engaging in the identified activities were collected during the baseline and comparison phases of the study to help assess whether or not the intervention resulted in socially important change that was noticeable to individuals who were not associated with the participants. Once all video was collected, the researchers selected representative video segments from the baseline and comparison sessions for each participant by determining the rate per minute of undesired behaviors for each video clip and then comparing that rate to the actual rate per minute of undesired behaviors for the entire observation session. One baseline phase and one comparison phase video clip for each target activity was selected for Adan, Brad and Ethan. One baseline phase and one comparison phase video was selected for one target activity for Daniel due to the fact that the bathroom area was the setting for one of Daniel's target activities.

Fourteen university students who were special education majors, were blind to the purpose of the study, and did not know the participants viewed the video segments that were presented in random order. After viewing each video segment, the university students completed a questionnaire developed by the researchers that utilized adapted Semantic Differential Scales (Osgood, Suci, & Tannebaum, 1957; Salcuni, DiRiso, Mazzeschi, & Lis, 2007) in order to measure the university students' perceptions of the study participants' behaviors. In completing the scale, raters indicated their perception of a behavior by marking a rating on a 7-point semantic scale that was defined by a pair of contrasting adjectives (Salcuni et al., 2007). For the purposes of this study, the adjective pairs on the semantic scale were *well behaved* and *poorly behaved*, *appropriate* and *inappropriate*, and *engaged* and *not engaged*.

Survey. A Likert scale survey designed by the researchers was used to examine the lead teachers' and classroom paraprofessionals' perceptions regarding: (a) the importance of the intervention strategies, (b) the usefulness of Social Stories as an instructional strategy, and (c) the format in which the Social Story was presented. The classroom teachers and paraprofessionals completed the survey for their respective participants during the final week of the study. The classroom teachers and paraprofessionals did not view data or graphs related to the intervention

outcomes prior to completing the survey, but they did observe at least one session of the baseline, comparison, flipped/most effective intervention, and maintenance phases of the study before completing the survey.

Reliability. Inter observer agreement was obtained to evaluate both procedural fidelity and dependent variable reliability across all phases of the investigation. To assess procedural fidelity, an independent observer watched and recorded the interventionist's implementation of a task-analyzed list of procedures. Procedural fidelity was collected during a minimum of 20% of both the comparison and the most effective/flipped phase sessions for all participants. Procedural fidelity was calculated by dividing the number of correct interventionist behaviors by the number of planned interventionist behaviors and multiplying the quotient by 100. Fidelity data showed that the interventionist correctly performed the planned interventions for 100% of the procedures for all participants.

To assess dependent variable reliability, an independent observer collected data related to the occurrence of the undesired behaviors during at least 20% of each of the baseline, comparison, most effective/flipped, and maintenance phase sessions for all participants. Next, the total number of target behaviors recorded per observation session across the two observers was compared and a percentage of dependent variable reliability was calculated. The mean dependent variability across participants for baseline was 98.75% (100% for Adan, 98% for Brad, 100% for Daniel, 97% for Ethan) and the mean dependent variability across participants for the comparison phase was 92.75% (89% for Adan, 94% for Brad, 88% for Daniel, 100% for Ethan). The mean dependent variable reliability for the most effective/flipped phase and the maintenance phase was 100% for all participants.

Data Analysis

Data regarding the frequency of undesired behaviors and the length (in minutes) of activities were used to compute the rate per minute of undesired behaviors for each activity. The rate per minute of undesired behaviors for each session was graphed to provide a visual representation of student behaviors across all phases of the investigation. Within-condition and between-condition visual analyses were conducted to examine data patterns related to variability, level, and trend (Lane & Gast, 2014).

In addition to within-condition and between-condition visual analysis, the researchers used the conservative dual-criterion method (CDC) to determine whether systematic changes in the behaviors occurred between different phases/conditions of the study (Fisher et al., 2003; Swobada, Kratochwill, & Levin, 2010). The CDC method calculates a statistical-based comparison between sets of data, and has been empirically validated as a method to improve the accuracy of the visual inspection of single-case data (Swoboda et al., 2010). The CDC method was used to compare (a) the baseline to comparison phases, and (b) the baseline to the most effective/flipped intervention phases for both the paper and iPad-based conditions for each participant. In order to determine whether a systematic change existed between the paper and iPad-based interventions, the CDC method was used to compare (a) the paper-based comparison condition to the iPad-based comparison condition, and (b) the paper-based most effective/flipped condition to the iPad-based most effective/flipped condition.

It is important to note that when using the CDC method, there is an order effect that influences the outcomes based upon which set of data is entered into the first column (e.g., baseline data) and which set of data is entered into the second column (e.g., intervention data). This order effect is appropriate when comparing baseline data to intervention data. However, this order effect presents a challenge when comparing two different interventions within the same condition (e.g., iPad-based comparison phase to paper-based comparison phase). As a result, when using the CDC method to analyze data within the same condition, the calculations were conducted in both directions and an a priori decision was made that a systematic change must be identified in both directions (i.e., paper to iPad and iPad to paper) when comparing interventions within a condition (i.e., within comparison phase, within flipped phase) in order to conclude that a systematic change between the two interventions existed.

Finally, in order to determine whether outcomes facilitated changes that were effective, efficient, and meaningful, the researchers summarized the social validity data obtained from (a) the Goal Attainment Scaling, (b) the teacher and paraprofessional rated-surveys, and (c) the video analysis of the participants' baseline and comparison phase behaviors. The completed video ratings were analyzed by calculating the mean score and standard deviation for the baseline and comparison phase ratings for each participant, as well as across all participants. One-tailed paired samples t-tests were conducted between the baseline and comparison phase mean scores in order to determine if the means of the baseline scores were significantly less than the means of the comparison phase scores.

Results

Results revealed that Social Stories presented in paper-based and iPad-based formats are effective in decreasing undesired behaviors when compared to baseline conditions. Furthermore, results suggest that a measurable difference did not exist between the efficiency and the effectiveness of the paper-based and iPad-based formats. Social validity assessment data suggests that both paper-based and iPad-based Social Story interventions are appropriate and effective interventions to decrease undesired behaviors.

The following sections summarize participant outcomes for each research question. Figures 1-4 illustrate the rate per minute of undesired behaviors per session across conditions for each participant. Data were examined for variability, level, and trend (Lane & Gast, 2014). The conservative dual-criterion method (CDC) was used to determine whether systematic changes in behaviors occurred between different phases of the study, and to compare the two different interventions within the same study phases (Swobada et al., 2010).

Effect of Paper-based Social Story Interventions

As illustrated in Figures 1-4, results revealed a decrease in rate per minute of undesired behaviors across all four participants following implementation of the paper-based Social Story intervention. The average rate per minute of undesired behaviors across all participants decreased by 72% (range=65%-82%) between the baseline and comparison phases, and by 81% (range=73%-97%) between the baseline and flipped intervention phases.

In baseline, Adan, Brad, & Ethan showed an increasing and variable trend and Daniel showed an increasing and stable trend of undesired behaviors in a contra-therapeutic direction. During the comparison phase, there was a decreasing and variable trend of undesired behaviors in a therapeutic direction for Adan, Brad and Ethan. Further, evaluations of behavior change from the baseline phase to the comparison phase for Adan, Brad, and Ethan revealed an immediate change in level, and the rate of undesired behaviors went from an increasing/accelerating trend in baseline to a decreasing/decelerating trend in the comparison phase. For Daniel, although a variable trend of undesired behaviors in a contra-therapeutic direction was noted between baseline and comparison, the absolute level change between baseline and comparison revealed a decreasing trend of undesired behaviors between the two conditions. The conservative dual-criterion (CDC) analysis indicated that systematic change occurred between the baseline and comparison phases for all participants.

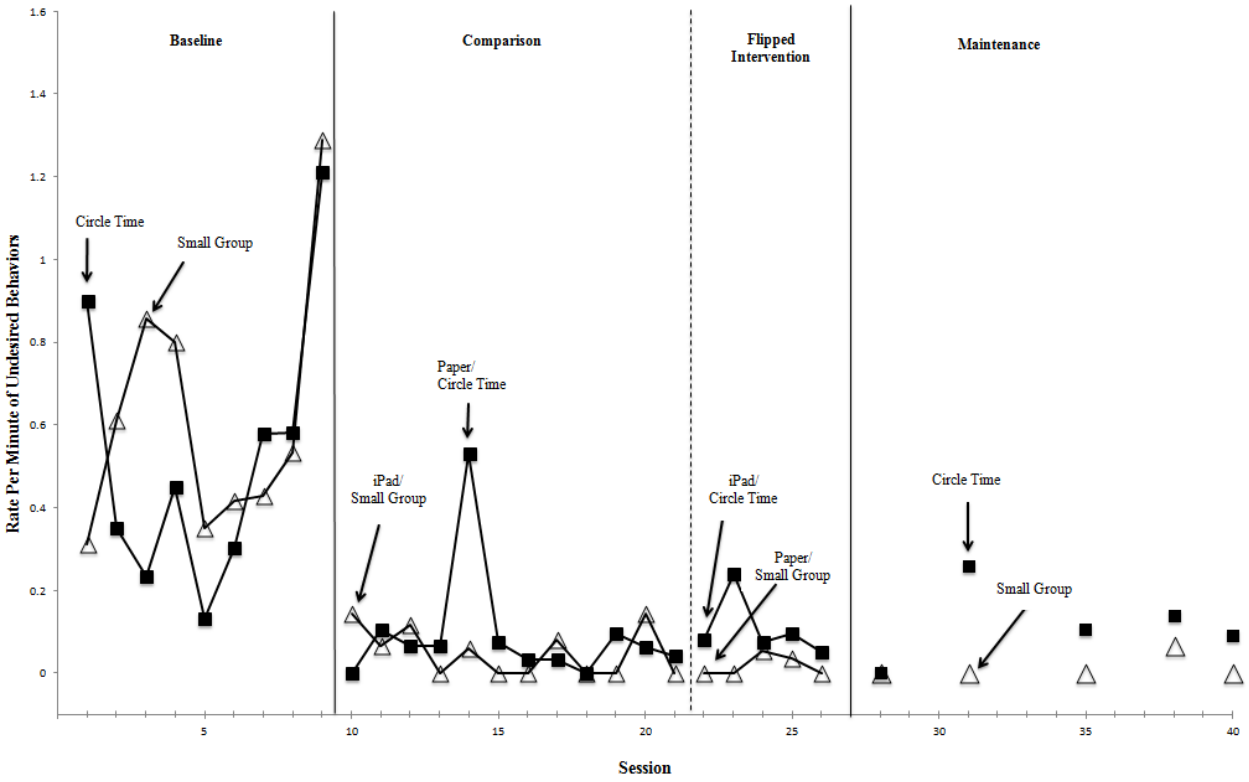

Figure 1. Line graph showing rates per minute of undesired behaviors during circle time and small group activities for Adan in baseline, comparison, flipped intervention, and maintenance phases.

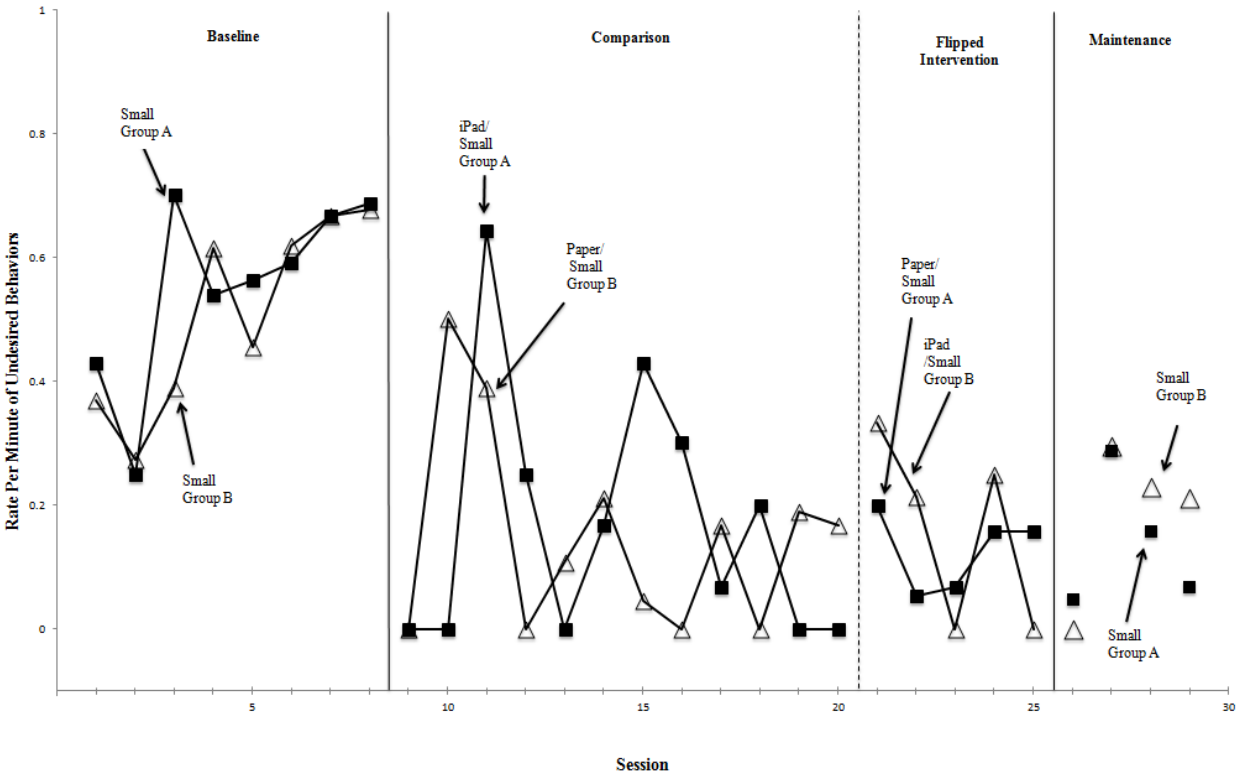

Figure 2. Line graph showing rates per minute of undesired behaviors during small group A and small group B activities for Brad in baseline, comparison, flipped intervention, and maintenance phases.

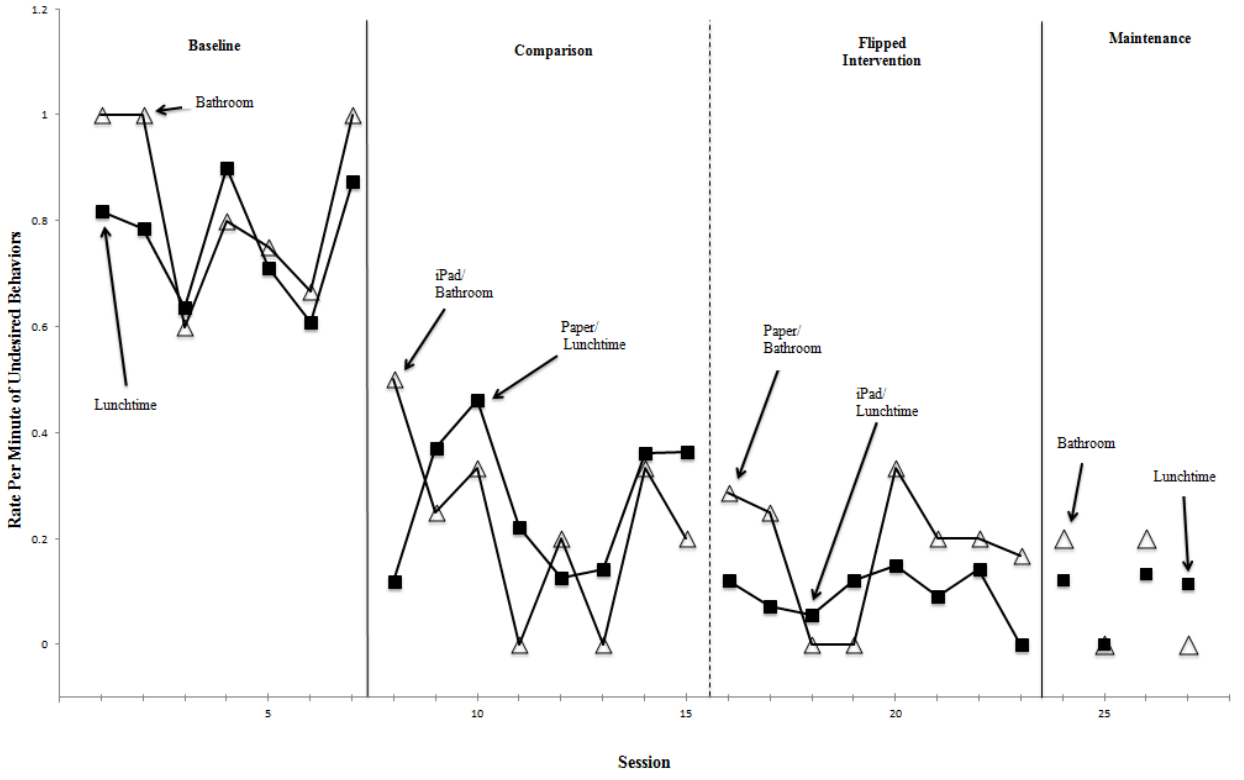

Figure 3. Line graph showing rates per minute of undesired behaviors during bathroom and lunchtime activities for Daniel in baseline, comparison, flipped intervention, and maintenance phases.

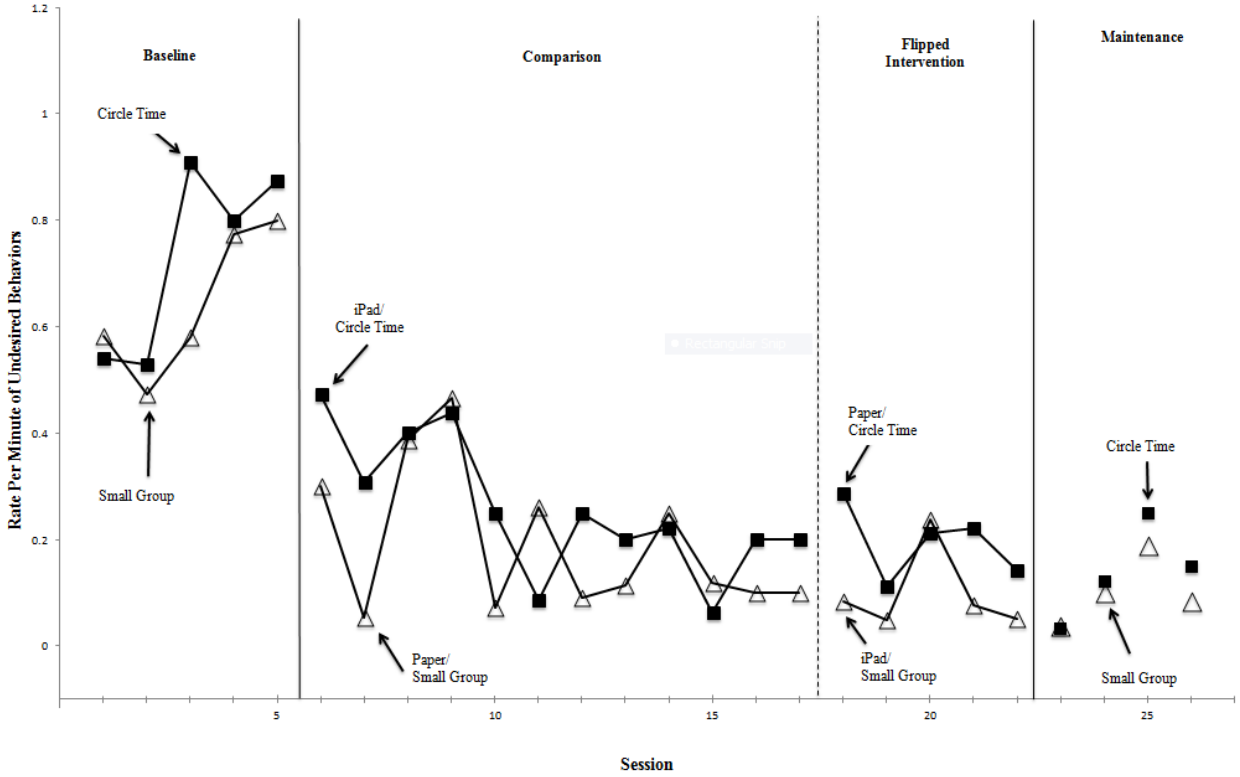

Figure 4. Line graph showing rates per minute of undesired behaviors during circle time and small group activities for Ethan in baseline, comparison, flipped intervention, and maintenance phases.

Given that a comparable level of performance was observed between the paper-based and iPad-based interventions during the comparison phase (see following sections), a flipped intervention phase (rather than a best intervention phase) was initiated for all participants. During the flipped intervention phase, all participants demonstrated a decreasing and variable trend of undesired behaviors in a therapeutic direction. Further, the CDC analysis indicated that systematic change occurred between baseline and flipped intervention phases for all participants.

During the maintenance phase of the study (a time in which participants had been exposed to both the paper-based and iPad-based interventions), the mean rate per minute of undesired behaviors for all participants was comparable to their means/ranges of undesired behaviors per minute during both their comparison and flipped intervention phases.

Effect of iPad-based Social Story Intervention

As illustrated in Figures 1-4, results revealed a decrease in rate per minute of undesired behaviors across all four participants following the implementation of the iPad-based Social Story intervention. The average rate per minute of undesired behaviors across all participants decreased by 75% (range=65%-92%) between the baseline and comparison phases, and by 81% (range=69%-88%) between the baseline and flipped intervention phases.

In baseline, Adan, Brad, and Ethan showed an increasing and variable trend of undesired behaviors in a contra-therapeutic direction and Danial showed a slightly decreasing (slope = -0.0184) and variable trend. During the comparison phase, a decreasing and variable trend of undesired behaviors in a therapeutic direction was noted for all participants. Evaluations of behavior change from the baseline phase to the comparison phase revealed an immediate change in level, and the rate of undesired behaviors went from an increasing/accelerating trend in baseline to a decreasing/decelerating trend in the comparison phase for Adan, Brad, and Ethan. For Daniel, the trend of undesired behaviors went from a slightly decreasing trend in baseline to a decreasing trend in the comparison phase and the absolute level change between both conditions for Daniel revealed a decreasing trend of undesired behaviors between the two conditions. The CDC analysis indicated that systematic change occurred between the baseline and comparison phases for all participants.

Given that a comparable level of performance was observed between the iPad-based and paper-based interventions during the comparison phase (see following sections), a flipped intervention phase was initiated for all participants. During the flipped intervention phase, all participants demonstrated a decreasing and variable trend of undesired behaviors in a therapeutic direction. Further, the CDC analysis indicated that systematic change occurred between baseline and flipped intervention phases for all participants.

During the maintenance phase of the study (a time in which all participants had been exposed to both the paper-based and iPad-based interventions), the mean rate per minute of undesired behaviors for all participants was comparable to their means/ranges of undesired behaviors per minute during both their comparison and flipped intervention phases.

Comparing the Effects of Paper-based and iPad-based Social Story Interventions

For each participant, a CDC analysis was conducted in order to assess whether there was a difference between the paper-based and iPad-based Social Story interventions during baseline, comparison, and flipped conditions. The CDC analysis for all participants indicated that systematic change did not exist between the two interventions in any of the conditions. It is notable that the CDC analysis for Daniel indicated a systematic change in one direction (comparing paper to iPad) but not in the reverse direction (comparing iPad to paper) in the flipped condition. However, as discussed previously, an a priori decision was made that a systematic change must be identified in both directions in order to conclude that a difference existed between the two interventions.

Social Validity

Social validity data was obtained through Goal Attainment Scaling (GAS), video ratings, and a survey. The following sections will summarize results.

Goal attainment scaling. The teachers' GAS ratings for each goal are provided in Table 2 and were analyzed using the GAS rating scale. As noted in Table 2, all participants achieved a GAS rating of "expected level of performance" (i.e., projected level of performance from the initiation of treatment and the behavior measurement period until the end of the behavior measurement period) or "better than expected level of performance" (i.e., somewhat more progress than expected during the treatment period) across activities.

Table 2

Goal Attainment Scaling Outcomes

Participant	Activity	Rating
Adan	Circle Time	0*
	Small Group	+1**
Brad	Small Group A	0
	Small Group B	0
Daniel	Lunchtime	+1
	Bathroom	+1
Ethan	Circle Time	0
	Small Group	0

Note. Adapted from "Goal Attainment Scaling: Applications, Theory and Measurement" by T.J. Kiresuk, A. Smith, and J.E. Cardillo, 1994, *Goal Attainment Scaling: Applications, Theory and Measurement.* Copyright 1994 by Erlbaum.

*0 = Projected/ expected level of performance

**+1= Better than expected level of performance

Video ratings. Based upon the responses of the 14 university students who completed the semantic differential scales for each of the video segments, mean scores and standard deviations were calculated. Table 3 summarizes the results of the video ratings for each participant as well as across all participants. As noted by Table 3, a significant difference was detected between the mean baseline score and the mean comparison phase score for each participant. Additionally, a significant difference was detected between the mean baseline scores and the mean comparison phase scores when the raw scores for all activities across all participants were combined.

Table 3
Social Validity Video Ratings

Participant	Activity	Baseline		Comparison		Significance (1-tailed)
		Mean	Standard Deviation	Mean	Standard Deviation	
Adan	Circle	37.333	6.506	81.333	6.658	0.01227*
	Small Group	28.333	1.528	72.0	13.0	0.01385*
Brad	Small Group A	25.0	4.583	88.0	3.606	0.00029*
	Small Group B	29.333	7.638	56.0	9.540	0.00217*
Daniel	Lunch-time	62.667	10.970	91.333	2.081	0.03133*
Ethan	Circle	56.333	5.132	86.667	4.933	0.00611*
	Small Group	63.0	14.799	80.667	8.144	0.02489*
All Participants	All Activities Combined	129.429	50.829	283.286	36.266	0.00035*

Note. *$p<.0$

Survey. At the conclusion of the study for each participant, the classroom lead teachers and paraprofessionals completed a survey regarding their perception of (a) the importance of the intervention strategies, (b) the usefulness of Social Stories as an instructional tool for students, and (c) the format in which the Social Story was presented. The mean result for the combined questions related to the importance of the Social Story intervention strategy across all participants for the paper-based Social Story interventions was 5.7 (range = 1-7; 1= strongly disagree, 7 = strongly agree) and 5.6 (range 1-7) for the iPad-based intervention. The majority of respondents (12 out of 16) felt that it would be easy to use the paper-based Social Story intervention and still meet the needs of the other children in the classroom. Similarly, the majority of respondents (14 out of 16) felt that it would be easy to use the iPad-based Social Story intervention and still meet the needs of the other children in the classroom. Finally, of the 16 survey respondents, 7 indicated that they did not have a preference between the two interventions**.** Of the respondents who had a preference between the paper-based and iPad-based interventions, 6 preferred the iPad and 3 preferred the paper.

Discussion

Results of this investigation revealed that Social Stories presented in both paper-based and iPad-based formats were effective in decreasing undesired behaviors when compared to baseline conditions and decreased rates of undesired behaviors continued during the maintenance phases of the study. The positive outcomes related to the impact of the paper-based Social Story intervention on rates of undesired behaviors add to the literature on the efficacy of paper-based Social Stories (e.g., Brownell, 2002; Chan & O'Reilly, 2008; Chan et al, 2011; Vandermeer, Milfor, Beamish, & Lang, 2013). The positive outcomes related to the impact of the iPad-based Social Story intervention on decreasing rates of undesired behaviors adds to the literature base examining the efficacy of computer and iPad based Social Stories (e.g., Chan et al., 2011;Vandermeer, Milford, Beamish, & Lang, 2013).

For all participants, results of CDC and visual analysis suggest that notable differences did not exist between the effectiveness and efficiency of the paper-based and iPad-based interventions. The results of this study do not support prior research suggesting that the effectiveness of Social Stories varies based on the story format (Brownell, 2002; Mancil et al., 2009). Plausible explanations for the difference in the results of the present study and previous studies may be due to issues related to the independent variables (e.g., variations in Social Story format) and/or differences in the experimental designs that were used across studies. Specifically, the current study used an adapted alternating treatment design (AATD) to examine the effects of two different treatments on two functionally independent, but equally difficult behaviors for each participant. In comparison, Brownell (2002) employed an ABAC/ACAB counterbalanced multiple-treatment design to examine the effects of two different treatments on one behavior, and Mancil et al. (2009) used an ABABCBC multicomponent reversal design to study the effects of two different treatments on one behavior.

Social Validity

This study used Goal Attainment Scaling (GAS), video ratings, and surveys to obtain social validity data. Obtaining data from more than one source and using more than one system of measurement was important in order to convincingly examine the social value of the intervention and its outcome (McDonnell & Tuesday Heathfield, 2011). The outcomes of the GAS, video ratings, and teacher/paraprofessional completed surveys provided strong support for the use of both paper-based and iPad-based Social Story interventions in early childhood classrooms. These results are similar to results from previous studies that have reported high social validity among educators with regard to Social Story interventions (e.g., Chan & O'Reilly, 2008; Ozdemir, 2008; Reynout & Carter, 2009).

Limitations

Given the wide range of abilities among young children who exhibit characteristics of ASD, as well as variability among early childhood classroom settings, it cannot be assumed that the results of this study would be replicated across other students and settings. Furthermore, data related to the rates of undesired behaviors were specific to each of the participant's target activities, and information related to the generalization of behaviors to other settings was not collected. The application and generalizability of the current investigation could be increased

through replications that utilize the present study methods with a wider range of participants, situations, interventionists, and settings.

Variability was noted in all phases of the investigation, and is a limitation of this study. Although the researchers controlled for several extraneous variables in order to decrease variability, future researchers may control for additional variables such as (a) ensuring that the same teacher is present for the target activities across time, and (b) ensuring that the sub-activities within activities (e.g., small group and circle) are consistent across time.

Another limitation of this study is that the investigator served as both the interventionist and data collector, and was aware of the purposes of the study. The investigator was used as the interventionist rather than a classroom teacher or paraprofessional in order to ensure consistency in the delivery of the intervention across participants. Although procedural fidelity and interobserver agreement measures helped control for potential researcher biases in this investigation, future studies could utilize different interventionists and/or independent data collectors who were blind to the purposes of the study.

Implications for Further Research

Recent systematic reviews and meta-analyses of Social Story research have revealed that variability exists with regard to the effectiveness of Social Story interventions (Kokina & Kern, 2010, Reynhout & Carter, 2011; Test, Richter, Knight, & Spooner, 2011). Given that Social Stories can be viewed as an intervention package that incorporates the use of several different strategies, it is likely that the varied effectiveness noted among Social Stories interventions will continue to persist if researchers do not investigate which strategy(ies) is/are contributing to positive outcomes. Specific strategies embedded within Social Story intervention packages may include; priming, behavioral momentum, prompting teachers/interventionists to focus on target behaviors, prompting students to focus on desired behaviors, visual supports, differential reinforcement, and shared book reading (see Johnston & Thompson, 2015). Future research examining the impact of one (or more) of these strategies on the effectiveness of Social Story interventions is warranted.

Implications for Practice

It is interesting to note that the social validity of Social Stories among educators is high (Reynhout & Carter, 2009). Given the high social validity among educators, it seems plausible to assume that practitioners will continue to utilize Social Story interventions. As a result, it is important for practitioners to increase the likelihood of effective and efficient results by considering issues related to patterns of change, format, and use of FBA data.

The results of the current study are in agreement with results from a meta-analysis conducted by Kokina and Kern (2010), suggesting that, if a Social Story is going to have an effect, the effect will likely occur rapidly following the introduction of the intervention. Thus, based on the results of this study and prior research, if an effect does not occur relatively quickly following the implementation of a Social Story intervention, a practitioner should consider an alternative intervention.

Results of the present study suggest that Social Story format (paper-based vs. iPad-based) does not have a measureable influence on behavioral outcomes. However, results of the present study's social validity survey questionnaire revealed that out of the sixteen respondents, 7 respondents did not have a preference for one format over another, 3 respondents preferred the paper-based format, and 6 respondents preferred the iPad-based format. These results suggest that, although behavioral outcomes may not vary across formats, interventionist preference may vary. Given this, practitioners should consider interventionist preference when choosing Social Story format.

Although Gray (2010) does not recommend a specific method for collecting information related to target situations and related behaviors, she does describe processes that are similar to the procedures of a functional behavioral assessment (FBA). The current study effectively utilized FBA strategies developed by O'Neill et al. (1997). Given the evidence provided in the present study, as well as existing evidence that supports the use of FBA in developing behavioral interventions (e.g., Horner et al., 2002), practitioners should utilize FBA data when developing Social Story interventions.

Conclusion

In conclusion, the results of this investigation revealed that both paper-based and iPad-based Social Story interventions were effective in decreasing rates of undesired behaviors. Furthermore, evidence suggested that the format of a Social Story did not make a significant difference in behavioral outcomes for the participants. In light of the fact that Social Story interventions are best conceptualized as an intervention package, future research should investigate which component(s) of Social Story interventions are contributing to their effectiveness. This, in turn, will inform the practices of interventionists who endeavor to utilize evidence-based practices.

References

Apple (2013). Apple Inc. (Version 5.3) [Mobile application software]. Retrieved from http://itunes.apple.com

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.

Chan, J. M., & O'Reilly, M.F. (2008). A social story intervention package for students with autism in inclusive classroom settings. *Journal of Applied Behavior Analysis, 41,* 405-409.

Chan, J.M., & O'Reilly, M. F., Lang, R.B., Boutot, E.M., White, P,J., Pierce, N., & Baker, S. (2011). Evaluation of a social stories intervention implemented by pre-service teachers for students with autism in general education settings. *Research in Autism Spectrum Disorders, 5*, 715-721.

Fisher, W.W., Kelley, M.E., & Lomas, J.E. (2003). Visual aids and structured criteria for improving visual inspection and interpretation of single-case designs. *Journal of Applied Behavior Analysis, 36,* 387-406.

Gray, C. (2004). Social stories 10.0: The new defining criteria & guidelines. *Jenison Autism Journal, 15*(4), 1-21.

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

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

Horner, R.H., Carr, E.G., Strain, P.S., Todd, A.W., & Reed, H.K. (2002). Problem behavior interventions for young children with autism: A research synthesis. *Journal of Autism and Developmental Disorders, 32,* 423-446.

Johnston, S. & Thompson, R. (2015). *What's working? An examination of the individual components of social story interventions*. Manuscript submitted for publication.

Kokina, A., & Kern, L. (2010). Social story interventions for students with autism spectrum disorders: A meta-analysis. *Journal of Autism and Developmental Disorders, 40,* 812-826.

Kuoch, H., & Mirenda, P. (2003). Social story interventions for young children with autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities, 18*(4), 219-227.

Lane, J.D., & Gast, D.L. (2014). Visual analysis in single case experimental design studies: Brief review and guidelines. *Neuropsychological Rehabilitation: An International Journal, 24*(3-4), 445-63.

Leaf, J., Oppenheim-Leaf, M., Leaf, R., Taubman, M., McEachin, J., Parker, T., Waks, A., Mountjoy, T. (2015). What is the proof? A methodological review of studies that have utilized social stories. *Education and Training in Autism and Developmental Disabilities, 50*(2), 127-141.

Mace, F., & Roberts, M. (1993). Factors affecting the selection of behavioral treatments. In J. Reichle, & D. Wacker (Eds.), *Communicative Approaches to the Management of Challenging Behavior* (pp. 113-133). Baltimore: Paul H. Brookes Publishing.

Mancil, G. R., Haydon, T., & Whitby, P. (2009). Differentiated effects of paper and computer-assisted social stories on inappropriate behavior in children with autism. *Focus on Autism and Other Developmental Disabilities, 24*(4), 205-215.

McDonnell, J.J., Jameson, M., & Rose, T. (2011). Alternating treatment designs. In R.E. O'Neill, J.J. McDonnell, F.F. Billingsley, & W. R. Jenson (Eds.), *Single case research designs in educational and community settings* (pp. 151-169). Upper Saddle River, NJ: Pearson Education.

McDonnell, J.J., & Tuesday Heathfield, L. (2011). Defining what to measure and how to measure it. In R.E. O'Neill, J.J. McDonnell, F.F. Billingsley, & W. R. Jenson (Eds.), *Single case research designs in educational and community settings* (pp. 16-38). Upper Saddle River, NJ: Pearson Education.

O'Neill, R.E., Horner, R.H., Albin, R.W., Sprague, J.R., Storey, K., & Newton, J.S. (1997). *Functional assessment and program development for problem behavior: A practical handbook* (2nd ed.). Pacific Grove, CA: Brooks/Cole.

Osgood,C.E., Suci, G.J., & Tannenbaum, P.H. (1957). *The measurement of meaning.* Urbana, Il: University of Illinois Press.

Ozdemir, S. (2008). The effectiveness of social stories on decreasing disruptive behaviors of children with autism: Three case studies. *Journal of Autism and Developmental Disorders, 38*(9), 1689-1696.

Reynhout, G., & Carter, M. (2009). The use of social stories by teachers and their perceived efficacy. *Research in Autism Spectrum Disorders, 3*(1), 232-251.

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

Roach, A.T., & Elliott, S.N. (2005). Goal attainment scaling: An efficient and effective approach to monitoring student progress. *Teaching Exceptional Children, 37*(4), 8-17.

Salcuni, S., DiRiso, D., Mazzeschi, S., & Lis, A. (2007). Parents representation of their children: An exploratory study using Osgood semantic differential scales. *Perceptual and Motor Skills, 105,* 39-46.

Swoboda, C. M., Kratochwill, T. R., & Levin, J. R. (2010). *Conservative dual-criterion method for single-case research: A guide for visual analysis of AB, ABAB, and multiple-baseline designs* (WCER Working Paper No. 2010-13). Retrieved from University of Wisconsin–Madison, Wisconsin Center for Education Research website: http://www.wcer.wisc.edu/publications/workingPapers/papers.php

Test, D. W., Richter, S., Knight, V., & Spooner, F. (2011). A comprehensive review and meta-analysis of the social stories literature. *Focus on Autism and Other Developmental Disabilities, 26*(1), 49-62.

Vandermeer, J., Milford, T., Beamish, W., & Lang, W. (2013). Using an iPad-presented social story to increase on-task behaviours of a young child with autism. *Neurodevelopmental Rehabilitation.* 146-154. doi: 10.3109/1718423.2013.809

About the Authors

Robyn M. Thompson, Ph.D., is an assistant professor and program director for the Occupational Therapy Assistant Program at Salt Lake Community College in Salt Lake City, UT. Robyn's research and scholarly interests include early childhood special education and supporting children with motor needs in inclusive settings.

Susan S. Johnston, Ph.D. is a Professor in the Department of Special Education at the University of Utah. Susan's research and scholarly interests include augmentative and alternative communication, early language and literacy intervention, and early childhood special education. She has published articles and has delivered state, national, and international presentations in these areas.