



## A Critical Review of Social Narratives

Justin B. Leaf<sup>1,2</sup> · Julia L. Ferguson<sup>1</sup> · Joseph H. Cihon<sup>1,2</sup> · Christine M. Milne<sup>1,2</sup> · Ronald Leaf<sup>1</sup> · John McEachin<sup>1</sup>

Published online: 16 July 2019

© Springer Science+Business Media, LLC, part of Springer Nature 2019

### Abstract

Social narratives, or story-based interventions, are defined as stories that describe social situations, appropriate social behaviors to display, and when to display the specified behaviors. Social narratives are a commonly implemented and empirically evaluated procedure used to improve social behavior and decrease the probability of aberrant behavior for individuals diagnosed with autism spectrum disorder. Although social narratives are a commonly implemented and evaluated procedure, recommendations about their use and effectiveness is conflicting. This paper reviews six interventions that fit the definition of social narratives (i.e., Social Stories™/social stories, social scripts, cartooning, comic strip conversations, power cards, and social autopsies). Fifteen articles were analyzed across multiple methodological dimensions to determine the level of evidence (i.e., convincing, partial, or not convincing). Results of the analysis indicated that the majority of social narrative studies did not demonstrate convincing evidence. Recommendations for clinicians and future research are discussed based on the results of the literature review.

**Keywords** Social narratives · Social stories · Story based · Story based intervention

Individuals diagnosed with autism spectrum disorder (ASD) and other neurodevelopmental disorders have a right to *effective* treatment (e.g., Van Houten et al. 1988). The effectiveness of behavior analytic interventions is also one of the dimensions of applied behavior analysis (Baer et al. 1968). Van Houten and colleagues (Van Houten et al. 1988) stated that “behavior analysts have an obligation to only use techniques that have been demonstrated by research to be effective, to acquaint consumers and the public with the advantages and disadvantages of these techniques, and to search continuously for the most optimal means of changing behavior” (p. 383).

---

✉ Justin B. Leaf  
Jblautpar@aol.com

<sup>1</sup> Autism Partnership Foundation, 200 Marina Dr., Seal Beach, CA 90740, USA

<sup>2</sup> Endicott College, Beverly, MA, USA

Furthermore, certifying boards have stated that it is the ethical responsibility of behavior analysts to have a reliance on scientific knowledge and provide intervention based upon scientific research (Behavior Analyst Certification Board 2014). Not only are behavior analysts obligated to use techniques backed by scientific research, but educators (Wong et al. 2013), speech pathologists (Dollaghan, 2007), and other service providers working with individuals diagnosed with ASD are also required to use practices that have been demonstrated to be effective in the scientific literature.

When determining what interventions to implement there are several criteria that should be evaluated. First, individuals should ensure that interventions are evidence-based, meaning they rely on clinical expertise, patient preference, and most importantly the *best empirical evidence* (National Autism Center 2015a). Furthermore, one should evaluate more stringent definitions of what is considered to be evidence-based (e.g., Cook and Cook 2011) to ensure the interventions they endorse, recommend, or implement meet the criteria to be evidence-based. Considering a more stringent criterion ensures interventionists are looking at all three prongs of evidence-based practice as opposed to just one or two. Second, one should ensure the procedures they implement have empirical support. This would mean that there have been empirical studies conducted on the interventions and/or the principles behind these interventions that support their effectiveness. The quality of empirical evidence can vary greatly and when evaluating the published, peer-reviewed research behind an intervention, individuals providing intervention for individuals diagnosed with ASD should look for important quality indicators including clear operational definitions, treatment fidelity data, and a functional relationship demonstrated through the correct implementation of the experimental design (Horner et al. 2005). Finally, the evidence found from any comparative research studies should be evaluated to ensure that the targeted intervention is the *most* effective procedure available.

To help determine which procedures are effective, evidence-based, and have empirical support, professionals have written commentaries (e.g., Horner et al. 2005), review papers (Wong et al. 2015), meta-analyses (Roth et al. 2014), and standards reports (e.g., National Autism Center 2015b). Additionally, researchers have conducted experimental analyses to evaluate the effectiveness of a variety of interventions. Sometimes these papers have resulted in universally accepted recommendations for clinicians. For example, reviews, commentaries, and standards projects have universally concluded that facilitated communication and rapid prompting method are not evidence-based and should not be implemented (e.g., National Autism Center 2015b). On other occasions, professionals have reached differing conclusions about the effectiveness of various interventions. For example, there have been several review papers which have warned against the implementation of Social Stories™ due to the lack of empirical evidence (e.g., Leaf et al. 2015; Reynhout and Carter 2011; Styles 2011). Yet, other commentaries and standards projects state that professionals could implement Social Stories™ as they meet the criteria to be evidence based (e.g., National Autism Center 2015b). When conflicting recommendations such as these exist, behavior analysts may be uncertain how to proceed.

Social narratives, also known as story-based interventions (referred to as social narratives throughout this paper), is an intervention that is implemented for individuals diagnosed with ASD (e.g., National Autism Center 2015b). Social narratives have been

defined as stories, represented visually, that describe various social situations, appropriate social behaviors to be displayed, and when these behaviors should be displayed (Zimmerman and Ledford 2017). Given this broad definition there are numerous interventions that meet this criterion including social scripts (Loveland and Tunali 1991), Social Stories™ (Gray and Garand 1993), cartooning (Coogle et al. 2017), Comic Strip Conversations™ (Gray 1994a), power cards (Campbell and Tincani 2011), and social autopsies (Bieber 1994). This broad definition has resulted in professionals providing different operational definitions for what constitutes a social narrative. In some instances, social narratives are used synonymously with Social Stories™ (e.g., Zimmerman and Ledford 2017), a trademarked intervention where professionals write a story following guidelines that have been created by Carol Gray (Gray 1994b). In other instances, social narratives are used to define non-trademarked social stories (i.e., stories that are written but that do not follow the guidelines laid out by Gray; e.g., Adams et al. 2004). In other instances, social narratives have also been used as a hypernym that includes all types of visually based stories to teach skills to individuals diagnosed with ASD (e.g., Social Stories™, comic strip conversations, and social autopsies; Wong et al. 2013). The lack of a clear operational definition may lead to confusion of what actually constitutes a social narrative.

For the purpose of this paper we define social narratives as a hypernym for six specific interventions: (1) Social Stories™/ social stories, (2) social scripts; (3) cartooning; (4) comic strip conversations; (5) power cards; and (6) social autopsies. Social narratives, as defined here, have been described as an evidence-based intervention by the National Autism Standards Project Phase 2 (National Autism Center 2015b) and by the National Professional Development Center on ASD (Wong et al. 2013). Yet other professionals have warned against the implementation of some of the interventions (e.g., Social Stories™) that fall under the hypernym of social narratives (e.g., Zimmerman and Ledford 2017). Furthermore, there may be little to no empirical research and/or reviews on other interventions that fall under the hypernym “social narratives” (e.g., comic strip conversations, social autopsies, and power cards).

Therefore, the purpose of this paper is to provide an evaluation and discussion of social narratives as they relate to interventions for individuals diagnosed with ASD and other neurodevelopmental disorders. In doing so, we describe each of the interventions that fall under the umbrella of social narratives. Next, we conduct a systematic search of the literature. Then we evaluate the studies found on quality indicators of experimental control (e.g., objective measurement, correct implementation of experimental designs) to classify each study as providing convincing evidence, partial convincing evidence, or not convincing evidence for the social narrative intervention. After experimental control has been evaluated and the articles have been classified, we then assess the quantity of empirical studies that were found to have convincing evidence to conclude if social narrative interventions would be considered to be evidence based. In this review, we include five of the six social narrative interventions (i.e., social scripts, cartooning, comic strip conversations, power cards, and social autopsies). It should be noted, we are not reviewing Social Stories™/social stories as there have been several reviews on that specific type of social narratives (e.g., Reynhout and Carter 2011; Styles 2011). While Social Stories™/social stories were not included in the literature review we do include them in our discussion of social narratives more generally.

## Social Narrative Interventions

### Social Scripts

Social scripts are a basic social narrative strategy in which the interventionist provides the learner with verbal statements or questions that they can use during social situations (Loveland and Tunali 1991). Within social scripts the interventionist typically has the learner practice using the script in contrived/analogue learning sessions as a way for the learner to become more fluent with the script. Once the learner displays the script appropriately in the analogue setting then the interventionist trains for generalization for the learner to display the skill in novel situations or with novel people.

### Cartooning

Cartooning is a form of a social narrative where the interventionist draws a cartoon to help explain the social situation and what behaviors the learner should engage in during different social situations (Coogle et al. 2017). Within cartooning, the interventionist also draws thought bubbles to display what the cartoon characters are feeling and thinking.

### Comic Strip Conversations™

Comic Strip Conversations™ are similar to cartooning in that the interventionist draws a series of cartoons to display the social situation and behaviors the learner should display. Comic Strip Conversations™ are based on the belief that visual supports help aid understanding and comprehension for individuals diagnosed with ASD. Comic Strip Conversations™ are an illustrative conversation between two or more people and are drawn out by at least two people together (Gray 1994a). Comic Strip Conversations™ use a set of eight symbols to denote basic conversation skills and also use specific colors to represent how others may feel about certain comments, thoughts, and questions used in the Comic Strip Conversation™ (Gray 1994a).

### Power Cards

Power cards are a type of social narrative that use a child's special interest or obsession as a motivator to teach the desired social behavior through a rule governed statement (Campbell and Tincani 2011). There are four steps when using the power card strategy. First, the child's interests and problem behaviors are determined through observation and interview with individuals that know the child well. Second, a functional behavior assessment is conducted to understand the contingencies surrounding the problem behavior. Third, the power card and scenario are developed based on the child's special interest, current comprehension level, and current reading level. The power card is then implemented with the child and the intervention is evaluated (Campbell and Tincani 2011).

## Social Autopsies

Social autopsies are one format of social narratives where the primary teaching modality is more vocal-verbal than visual. Social autopsies occur after the learner engages in an incorrect social behavior. At this point the interventionist has the learner identify the incorrect social behavior, who was harmed by the incorrect social behavior, how to correct the social mistake, and develop a plan to ensure that it does not occur in the future (Bieber 1994).

## Methods

### Inclusion Criterion

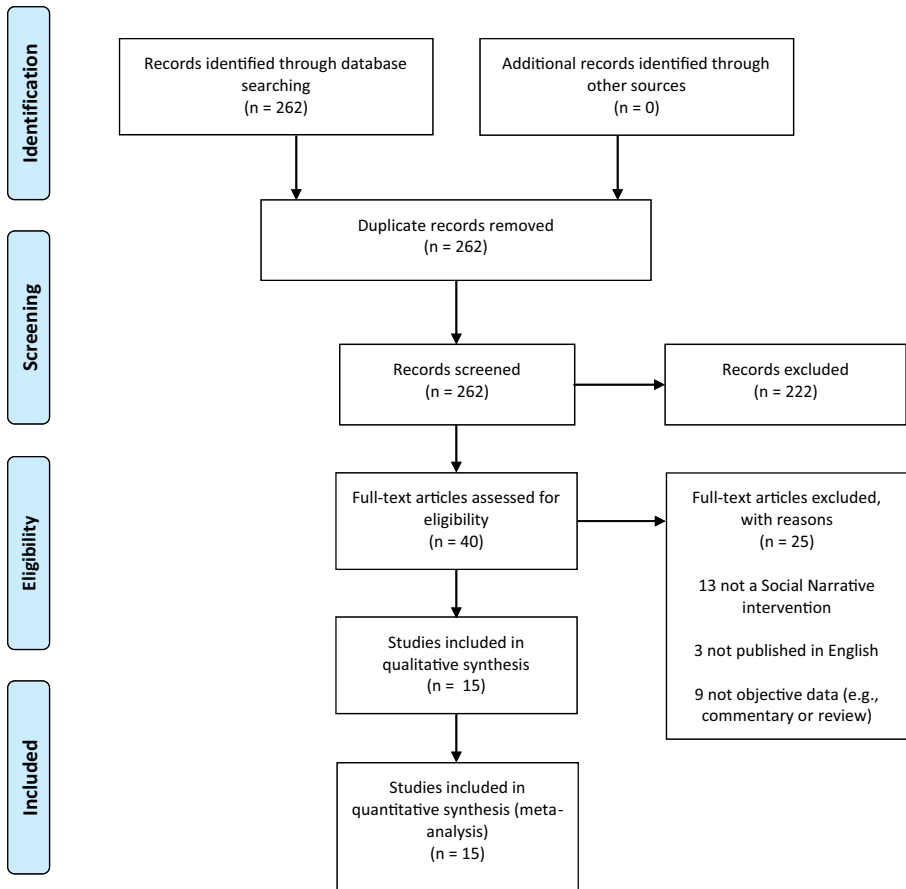
To be included in this review an article had to meet the following criteria. First, articles had to be published in a peer reviewed journal between January 1950 and November 2018. Second, the studies had to have either social scripts, cartooning, comic strip conversations, power cards, or social autopsies as the independent variable. Third, the study had to include at least one participant diagnosed with ASD, DD, or ID. Fourth, articles had to be accessible either through a university online library, or available for purchase through the journal. Finally, articles had to be written or available in English.

### Search Procedure

We conducted a systematic search of the social narratives in accordance with the PRISMA guidelines (Moher et al. 2009; see Fig. 1). The PRISMA guidelines consists of four broad levels: (1) identification of articles, (2) screening of articles; (3) eligibility of articles, and (4) inclusion of articles. Two reviewers (referred to as primary and secondary reviewer hence forward) were used across all levels.

The primary database that was searched was PsychINFO from the years of January 1950 to November 2018. The search terms were: ["power cards" OR "social autopsies" OR "comic strip conversations" OR "cartooning" OR "social articles" OR "social narratives OR "story based interventions" OR "social scripts"] all paired with "autism spectrum disorder" or "developmental disabilities." PsychINFO search yielded 262 articles. Two reviewers read the title and abstract of the 262 articles and retrieved the full text of any article ( $n = 40$ ) which appeared to implement one of the social narrative interventions. Interrater agreement was taken for which articles were going to be excluded and which articles were going to be assessed for eligibility. To conduct interrater agreement, we calculated the number of agreements (articles chosen to assess eligibility or excluded) over the total number of articles screened. Interrater agreement for this measure was 100%.

Next, two reviewers evaluated the 40 articles to identify if the studies met the inclusion criterion stated above. A total of 15 articles were deemed relevant; while 25 were deemed not to meet the inclusion criterion. Interrater agreement was taken on which articles were going to be excluded and which articles were going to be included for further analysis. For interrater agreement we calculated the number of agreements



**Fig. 1** PRISMA diagram

(articles included or excluded) over the total number of articles screened. Interrater agreement on this measure was 100%.

## Measures

**Experimental Control** Our first measure was to evaluate the articles to assess if the researchers implemented an experimental design that met previously described standards (Campbell and Stanley 1963; Kazdin 2011; Kennedy 2005). To evaluate experimental designs, we modified an evaluation of experimental designs previously implemented by Leaf and colleagues (Leaf et al. 2015). That is, we used specific criteria developed for each type of design that implemented in the research articles (i.e., AB case design, reversal design, multiple baseline design, group design). The criteria used to evaluate each study was based on the research design implemented. This evaluation method was chosen because this method analyzes the experimental rigor of the design implemented in more ways than other evaluation methods have done. For example, the National Autism Center (2015a, b) used the Scientific Merit Rating Scale to evaluate research articles for interventions implemented with individuals with ASD. Although

the Scientific Merit Rating Scale takes the research design into account (i.e., group design, single subject designs except alternating treatment design, and alternating treatment design), they only consider some important variables when analyzing a research design. For example, when analyzing research using a single-subject design other than an alternating treatment design (e.g., reversal design) a rating of 5 would be received if the study had a minimum of three comparisons of intervention and control conditions, if the number of data points per condition was greater than 5, the number of participants was more than three, and there was no data loss. Although these are critical variables to a single-subject research design, it does not evaluate other important aspects to a correctly implemented research design such as ensuring baseline data is trending in the correct way prior to implementing the intervention, demonstrating clear behavior change (i.e., non-overlapping data) in the intervention condition, and the immediacy of the effect of intervention. Below are the criteria and definitions used to evaluate each research design.

**Case Study** Table 1 displays the criteria used to evaluate and classify studies that implemented a case study design. First, the type of data collected was evaluated (i.e., objective, subjective). Next, the length of baseline was evaluated. The baseline trend was also evaluated. Stability in a baseline trend was defined as (a) two consecutive sessions trending in the right direction, or (b) two consecutive days of stability, or (c) three out of four sessions at a stable rate without the last data point heading in the wrong direction, and (d) criteria had to be met across all applicable skills and/or participants. Next, the immediacy of the effect of intervention was evaluated. Immediacy of effect included evaluating if the behavior change occurred within three sessions or after three sessions and also included an evaluation of (a) if two of the three data points were higher (for increasing behaviors) or lower (for decreasing behaviors) than all data points during baseline or (b) if the third data point of intervention was higher than all data points during baseline. Overlapping data was also assessed through the percentage of sessions where an intervention data point was the same as any baseline

**Table 1** Criteria for studies using a case study design

Level of Demonstration	Type of Data	Length of Baseline	Baseline Trending	Effect Immediate	Overlapping Data	Other Procedures
Convincing	Objective	3 or more sessions	Stable or trending in correct direction	Behavior change within 3 sessions	20–0% overlapping data (ALL)	Not Combined with other procedures
Partial	Objective	1 to 3 sessions	Stable or trending in correct direction	Behavior change within 3 sessions	40–21% overlapping data (ALL)	Combined with other Procedures
Not Convincing	Subjective	0 sessions or not reported	No stability or not trending in the correct direction	Behavior change occurring after 3 sessions	100 to 41% overlapping data (ALL)	Combined with other Procedures

level data points, Finally, case study designs were evaluated to see if the social narrative intervention was combined with any other procedures.

**Reversal** Table 2 displays the criteria used to evaluate and classify studies that implemented a reversal design. First, the type of data collected was evaluated (i.e., objective, subjective). Next, the baseline trend was evaluated prior to implementing intervention. This was defined as (a) two consecutive sessions trending in the right direction, or (b) two consecutive sessions of stability, or (c) three out of four sessions at a stable rate without the last data point heading in the wrong direction, and (d) criteria had to be met across all participants and/or skills. The intervention condition trend was also assessed based on this criteria: (a) the last two data points trending in the correct direction and higher than 85% of all baseline points, or (b) the last two data points are stable and higher than 85% of baseline sessions, or (c) three out of four sessions are at a stable rate, are higher than 85% of all baseline sessions, without the last data point heading in the wrong directions, and (d) criteria was met across all skills and/or participants. Next, the behavior change was assessed. This was defined as (a) 75% of all data points in a condition are higher (for increasing behaviors) or lower (for decreasing behaviors) than all of the baseline data points or (b) clear level change through visual analysis. Finally, each study that implemented a reversal design was assessed to see if the social narrative intervention was combined with any other procedures.

**Multiple Baseline Design** Table 3 displays the criteria used to evaluate and classify studies that implemented a multiple baseline design. First, the type of data collected was evaluated (i.e., objective, subjective). Next the baseline trend prior to intervention was evaluated. This was defined as (a) two consecutive sessions trending in the right direction, or (b) two consecutive sessions of stability, or (c) three out of four sessions at a stable rate without the last data point heading in the wrong direction, and (d) criteria had to be met across all participants and/or skills. Next the staggering of the intervention was evaluated. The definition for staggering correctly was defined as (a) trending correct was on the panel directly above without the previous two data points trending

**Table 2** Criteria for studies using a reversal design

Level of Demonstration	Type of Data	All Baseline Trending the Correct Way Prior to Reversal	All Intervention Trending the Correct Way Prior to Reversal	Clear Behavior Change	Other Procedures
Convincing	Objective	100% of all baseline condition	100% of all intervention conditions	100 to 80% of all intervention conditions	Not Combined with other Procedures
Partial	Objective	99 to 50% of all baseline conditions	99 to 50% of all intervention conditions	79 to 50% of all intervention conditions	Combined with other procedures
Not Convincing	Subjective	49% to 0% of all baseline conditions	49 to 0% of all baseline conditions	49 to 0% of all intervention conditions	Combined with other procedures



**Table 3** Criteria for studies using a multiple baseline design

Level of Demonstration	Type of Data	All baseline trending the correct way prior to intervention	Staggering Correct	Clear Behavior Change	Other Procedures	Number of Legs
Convincing	Objective	100% of all stable	100 to 75% correctly staggered	100 to 80% of all conditions	Not Combined with other Procedures	3 or More Legs
Partial	Objective	99 to 67% stable	74 to 50% correctly staggered	79 to 50% of all conditions	Combined with other procedures	2 Legs
Not Convincing	Subjective	66 to 0% stable	49 to 0% correctly staggered	49 to 0% of all conditions	Combined with other procedures	Less than 2 legs

the incorrect direction and (b) stability on the last two data points were higher than 80% of the baseline data points. Next, the behavior change was assessed. This was defined as (a) 75% of all data points in a condition are higher (for increasing behaviors) or lower (for decreasing behaviors) than all of the baseline data points or (b) clear level change through visual analysis. Each study that implemented a multiple baseline design was also assessed to see if the social narrative intervention was combined with any other procedures. Finally, studies using a multiple baseline design were evaluated on the number of legs implemented in the multiple baseline design.

**Group Design** Table 4 displays the criteria used to evaluate and classify studies that implemented a group design. First the type of data collected was evaluated (i.e.,

**Table 4** Criteria for studies using a group design

Level of Demonstration	Type of Data	Groups	Randomized	Evaluator	Pre-Test	Post Test
Convincing	Objective	Control Group or 2nd Treatment Group	Randomized, Quasi Randomized, or Match to Sample	Blind Evaluator	Occurred Both Groups	Occurred Both Groups
Partial	Standardized Assessment	N/A	N/A	Research or Teacher	Occurred with Treatment Group	Occurred with Treatment Group
Not Convincing	Subjective	No Second Group	No Randomized, Quasi Randomized, or Match to Sample	Child or Caregiver	No Pre-test	No Post-Test

objective, standardized assessment, subjective). Next the study was evaluated for the groups included (i.e., control group or 2nd treatment group, no second group). The randomization of the groups was also evaluated (i.e., randomized, quasi randomized, or match to sample vs. no randomization, quasi randomization, or match to sample). Studies implementing a group design were also assessed on the type of evaluator they used and whether it was a blind evaluator, a researcher or teacher, or a child or caregiver. Finally, each study that implemented a group design was evaluated for if a pre-test and post-test occurred for both groups, if a pre-test and post-test only occurred with the treatment group, or no pre-test and post-test occurred.

**Level of Demonstration** Based on the criteria described above, a study was determined as either having convincing evidence, partial evidence, or not convincing evidence. The definitions for convincing evidence, partial evidence, and not convincing evidence for each type of research design are displayed in Tables 1, 2, 3, and 4. A study was classified as convincing, partially convincing, or non-convincing based upon their lowest score across the variables evaluated. Interrater agreement was conducted on 40% of articles evaluated for study classification. Interrater agreement was 100% for experimental control study classification.

**Quantity of Evidence Base** Once the quality of evidence was assessed, our second measure was to evaluate if the studies that had convincing evidence met the basic quantity requirement to establish the intervention as an evidence-based practice laid out by Horner et al. (2005). Specifically, Horner and colleagues' criteria included a minimum of five experimental studies published in peer-reviewed journals, conducted by at least three separate research groups, and including at least 20 participants collectively across the studies. Thus, we evaluated the number of studies that were found to have convincing evidence that were conducted for each social narrative intervention published in a peer-reviewed journal. Second, we evaluated the number of different research groups that conducted the convincing research studies on the interventions and if they met the minimal threshold of at least three separate research groups conducting the research studies. Finally, we evaluated the number of participants per convincing study and the total number of participants across all convincing studies for each type of social narrative intervention.

## Results

Table 5 provides an overview of the 15 studies which were evaluated in this review and denotes the type of social narrative intervention evaluated, the research design used, the quality rating, and why that quality rating was given to each study. If studies were found to have convincing evidence then the total number of studies found to be convincing, the participants in each study, and the number of research groups conducting the research was evaluated.

**Table 5** Overview of studies

Authors	Intervention	Number of Participants	Research Design	Level of Convincing Evidence	Why Designation of Evidence
Ahmed-Husain and Dunsmuir (2014)	Comic Strip Conversations™	8	Multiple baseline	Not convincing	Trending baseline levels
Angell et al. (2011)	Power cards	3	Reversal	Not convincing	Trending baseline levels
Campbell and Tincani (2011)	Power cards	3	Multiple baseline	Partial convincing evidence	Trending baseline levels
Daubert et al. (2015)	Power cards	2	Multiple baseline	Partial convincing evidence	Trending baseline levels
Davis et al. (2010)	Power cards	3	Multiple baseline	Convincing evidence	N/A
Ganz et al. (2008)	Social scripts	3	Multiple baseline	Not convincing	Trending baseline levels
Hundert et al. (2014)	Social scripts	3	Multiple baseline	Partial convincing evidence	Staggering correctly & other procedures
Hutchins and Prelock (2006)	Comic Strip Conversations™	2	Case study	Not convincing	Type of data, immediate effect, & overlapping data.
Hutchins and Prelock (2008)	Comic Strip Conversations™	1	Reversal	Not convincing	Type of data, clear behavior change,
Hutchins and Prelock (2012)	Comic Strip Conversations™	17	Multiple baseline	Not convincing	Type of data & staggering correctly
Keeling et al. (2003)	Power cards	1	Multiple baseline	Not convincing	Trending baseline levels
Loveland and Tunali (1991)	Social scripts	13	Group design	Partial convincing evidence	Evaluator
Parker and Kamps (2011)	Social scripts	2	Multiple baseline	Partial convincing evidence	Trending baseline levels
Pierson and Glaeser (2005)	Comic Strip Conversations™	4	Case study	Not convincing	Type of data, & data not reported
Pierson and Glaeser (2007)	Comic Strip Conversations™	3	Case study	Not convincing	Length of baseline, baseline trend, no immediate effect

**Social Scripts** There was a total of four publications on social scripts published by four different research groups found through the literature search. When evaluating the four social script studies with respect to experimental control, three of the studies demonstrated partially convincing evidence (i.e., Hundert et al. 2014; Loveland and Tunali 1991; Parker and Kamps 2011) while one study showed no convincing evidence (i.e., Ganz et al. 2008). Since none of the social script studies were found to have a convincing level of evidence, the quantity of studies that could contribute to the establishment of an evidence-based practice could not be assessed. Therefore, there were zero social scripts studies to meet the quantity standards for an evidence-based practice according to Horner and colleagues (Horner et al. 2005).

**Comic Strip Conversations™** A total of six publications were found from the literature search on Comic Strip Conversations™. When evaluating the six publications with respect to experimental control (i.e., Leaf et al. 2015), all six of the studies demonstrated no convincing evidence (i.e., Ahmed-Husain and Dunsmuir 2014; Hutchins and Prelock 2006, 2008, 2012; Pierson and Glaeser 2005, 2007). Since none of the studies evaluated demonstrated convincing evidence, Comic Strip Conversations™ did not meet the quantity standards to be considered an evidence-based practice according to Horner et al. (2005).

**Power Cards** There were a total of five publications found on power cards. When evaluating the five publications with respect to experimental control, only one study was found to demonstrate a convincing level of evidence (i.e., Davis et al. 2010), two studies were found to have partial convincing evidence (i.e., Campbell and Tincani 2011; Daubert et al. 2015), and two were found to have no convincing evidence (i.e., Angell et al. 2011; Keeling et al. 2003) that power cards were responsible for the behavior change. The one publication that was found to have convincing evidence was conducted by one research group and had a total of 3 participants. Thus, there was an insufficient number of power card studies to meet the quantity standards for an evidence-based practice according to Horner et al. (2005).

**Social Autopsies** There were no peer-reviewed empirical studies that evaluated social autopsies on changing behavior for individuals diagnosed with ASD, DD, or ID. Since we found no peer-reviewed published studies it was impossible to determine the quality of experimental control as outlined by Leaf et al. 2015. Therefore, social autopsies did not meet the quantity standards to be considered an evidence-based practice according to Horner et al. (2005).

**Cartooning** There were no peer-reviewed studies published evaluating cartooning with individuals diagnosed with ASD, DD, or ID. Since we found no peer-reviewed published studies it was impossible to determine the quality of experimental control as outlined by Leaf et al. 2015. Thus, cartooning did not meet the minimum quantity standards to be considered evidence-based practice according to Horner et al. (2005).

## Discussion

The purpose of this review was to evaluate procedures that are commonly defined under the hypernym of social narratives. While conducting the review we assessed if the research conducted demonstrated convincing levels of experimental control between the intervention implemented and the corresponding change in behavior as defined by Leaf et al. (2015). We also assessed if the studies found to have convincing levels of evidence resulted in the minimum quantity standards to be considered an evidence-based practice as defined by Horner et al. (2005). A total of 15 studies were found that evaluated social narratives. When evaluating if sufficient experimental control was established across the 15 studies, only 1 study was determined to show convincing evidence, 5 studies (i.e., 33.3%) were determined to have partially convincing evidence, and 9 studies (i.e., 60%) were determined to show no convincing evidence. Since only one study demonstrated convincing evidence, none of the interventions under the hypernym of social narratives (i.e., social scripts, Comic Strip Conversations™, power cards, social autopsies, cartooning) met the quantity standards to be considered an evidence-based practice as defined by Horner et al. (2005). Therefore, the results show the interventions under the hypernym “social narratives” do not meet the basic quantity criteria to be considered an evidence-based practice and the majority of studies did not demonstrate a convincing functional relationship between the intervention and corresponding change in behavior.

We did not evaluate Social Stories, a type of social narrative that is widely used, because there have been numerous reviews that have already been written (e.g., Leaf et al. 2015; Reynhout and Carter 2011; Rhodes 2014; Styles 2011). Although there is sufficient research on Social Stories to meet the quantity criteria defined by Horner et al. (2005), previous reviews have shown that the majority of studies only showed partially convincing evidence (41.5% of studies) or no convincing evidence (51.2% of studies; Leaf et al. 2015).

This review also has implications for interventionists, researchers, and entities using and producing evidence-based standards. Interventionists have a legal, ethical (e.g., Behavior Analyst Certification Board 2014), and moral obligation to implement only interventions that are considered to be evidence-based and which have been shown to be effective for the clients whom they serve, especially when effective alternatives are available. Failure to do so could result in a waste of time, money, and could be emotionally taxing for families to spend valuable time on non-evidence-based interventions. Interventionists should avoid the use of social narratives given the lack of empirical research documenting that the interventions are responsible for observed changes. In that vein, interventionists should implement interventions that meet the minimum quality and quantity standards to be considered an evidence-based practice and convincing levels of experimental control (e.g., video modeling, discrete trial teaching, behavioral skills training).

Although, the current research has not convincingly shown that social narrative interventions are responsible for behavior change, researchers should continue to evaluate social narrative interventions. In doing so, researchers should first and

foremost ensure appropriate experimental rigor by ensuring the research design is employed correctly to demonstrate a functional relationship between the independent and dependent variable. Single-subject research methodology may be best suited to evaluate the conditions under which social narrative interventions are and are not effective. Individual differences could be clearly identified and further evaluated. If the use of single-subject designs are effective at identifying certain conditions under which social narratives are effective, researchers could begin to conduct comparative studies between social narrative interventions and other established interventions (e.g., video modeling). The results of this line of research could be greatly beneficial for interventionists designing and implementing interventions. That is, this research could inform when and when not to use social narrative interventions.

Finally, this review and previous reviews also have implications for entities that provide standards of practice (e.g., National Autism Center 2015b). First, we encourage that these entities not only evaluate minimum quantity standards that constitute an evidence-based practice (e.g., number of participants, number of studies) but also evaluate if these studies have demonstrated a convincing functional relationship between the independent and dependent variables by evaluating the trends in the data during baseline and intervention, the immediacy of the effect, and if clear behavior change was demonstrated. Second, while we understand the potential rationales for using hypernyms when defining evidenced or best practices, there are potential challenges. For instance, lacking clear definitions, criteria, or parameters for which procedures or interventions fall within the hypernym may result in interventionists finding it difficult to distinguish which procedures or interventions constitute the evidence base. This could also lead to a misunderstanding about whether a procedure is or is not considered to be evidence based. If hypernyms are unavoidable, our recommendation is to provide clear definitions, criteria, and parameters for the intervention or procedure and to update them frequently.

We conducted this review on interventions that have been considered to be social narrative interventions. While this review focused on social narratives, it is possible that several other perceived evidence-based practices could be evaluated in a similar vein. For instance, “Social Skills Package” is a common term to define a broad range of intervention packages. It is possible that there are interventions that consider themselves or are marketed as a “Social Skills Package” with varying levels of convincing evidence of their effectiveness. It is our hope that researchers continue to evaluate these types of interventions, and that clinicians utilize critical thinking when evaluating the evidence base for any purported established intervention.

## Compliance with Ethical Standards

**Ethical Approval** This article does not contain any studies with human or animal participants performed by any of the authors.

**Informed Consent** As such no informed consent was needed in this study.

**Conflict of Interest** None of the authors have any conflict of interests with the information presented within this article.

## References

- Adams, L., Gouvousis, A., VanLue, M., & Wladron, C. (2004). Social story intervention: Improving communication skills in a child with autism spectrum disorder. *Focus on Autism and Other Developmental Disabilities, 19*(2), 87–94.
- Ahmed-Husain, S., & Dunsmuir, S. (2014). An evaluation of the effectiveness of comic strip conversations in promoting the inclusion of young people with autism spectrum disorder in secondary schools. *International Journal of Developmental Disabilities, 60*(2), 89–108.
- Angell, M. E., Nicholson, J. K., Watts, E. H., & Blum, C. (2011). Using a multicomponent adapted power card strategy to decrease latency during interactivity transitions for three children with developmental disabilities. *Focus on Autism and Other Developmental Disabilities, 26*(4), 206–217.
- Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis, 1*(1), 91–97.
- Behavior Analyst Certification Board. (2014). *Professional and ethical compliance code for behavior analysts*. Littleton: Author.
- Bieber, J. (1994). *Learning disabilities and social skills with Richard LaVoie: Last one picked...First one picked on*. Washington, DC: Public Broadcasting Service.
- Campbell, D. T., & Stanley, J. (1963). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.
- Campbell, A., & Tincani, M. (2011). The power card strategy: Strength-based intervention to increase direction following of children with autism spectrum disorder. *Journal of Positive Behavior Interventions, 13*(4), 240–249.
- Coogle, C. G., Ahmed, S., Aljaffal, M. A., Alsheef, M. Y., & Hamdi, H. A. (2017). Social narrative strategies to support children with autism spectrum disorder. *Early Childhood Education Journal, 46*, 445–450. <https://doi.org/10.1007/s10643-017-0873-7>.
- Cook, B. G., & Cook, S. C. (2011). Unraveling evidence-based practices in special education. *The Journal of Special Education, 47*(2), 71–82.
- Daubert, A., Hornstein, S., & Tincani, M. (2015). Effects of a modified power card strategy on turn taking and social commenting of children with autism spectrum disorder playing board games. *Journal of Developmental and Physical Disabilities, 27*, 93–110.
- Davis, K. M., Boon, R. T., Cihak, D. F., & Fore, C. (2010). Power cards to improve conversational skills in adolescents with Asperger syndrome. *Focus on Autism and Other Developmental Disabilities, 25*(1), 12–22.
- Dollaghan, C. A. (2007). *The handbook for evidence-based practice in communication disorders*. Baltimore, MD: Brookes.
- Ganz, J. B., Kaylor, M., Bourgeois, B., & Hadden, K. (2008). The impact of social scripts and visual cues on verbal communication in three children with autism spectrum disorder. *Focus on Autism and Other Developmental Disabilities, 23*(2), 79–94.
- Gray, C. (1994a). *Comic strip conversations: Illustrated interactions that teach conversation skills to students with autism and related disorders*. Jenison: Jenison Public Schools.
- Gray, C. (1994b). *Social stories*. Arlington: Future Horizons.
- Gray, C. A., & Garand, J. D. (1993). Social stories: Improving responses of students with autism with accurate social information. *Focus on Autistic Behavior, 8*, 1–10.
- Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children, 71*(2), 165–179.
- Hundert, J., Rowe, S., & Harrison, E. (2014). The combined effects of social script training and peer buddies on generalized peer interaction of children with ASD in inclusive classrooms. *Focus on Autism and Other Developmental Disabilities, 29*(4), 206–215.
- Hutchins, T. L., & Prelock, P. A. (2006). Using social stories and comic strip conversations to promote socially valid outcomes for children with autism. *Seminars in Speech and Language, 27*(1), 47–59.
- Hutchins, T. L., & Prelock, P. A. (2008). Supporting theory of mind development: Considerations and recommendations for professionals providing services to individuals with autism spectrum disorder. *Topics in Language Disorders, 28*(4), 340–364.



- Hutchins, T. L., & Prelock, P. A. (2012). Parents' perspective of their children's social behavior: The social validity of social stories™ and comic strip conversations. *Journal of Positive Behavior Interventions*, *15*(3), 156–168.
- Kazdin, A. E. (2011). *Single-case research designs: Methods for clinical and applied settings* (2nd ed.). New York: Oxford University Press.
- Keeling, K., Myles, B. S., Gagnon, E., & Simpson, R. L. (2003). Using the power card strategy to teach sportsmanship skills to a child with autism. *Focus on Autism and Other Developmental Disabilities*, *18*(2), 105–111.
- Kennedy, C. H. (2005). *Single-case designs for educational research*. Boston, MA: Pearson Education, Inc..
- Leaf, J. B., Oppenheim-Leaf, M. L., Leaf, R. B., Taubman, M., McEachin, J., Parker, T., Waks, A. B., & 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.
- Loveland, K. A., & Tunali, B. (1991). Social scripts for conversational interactions in autism and down syndrome. *Journal of Autism and Developmental Disorders*, *21*(2), 177–186.
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, *6*(7), e1000097. <https://doi.org/10.1371/journal.pmed.1000097>.
- National Autism Center. (2015a). *Evidence-based practice and autism in the schools* (2nd ed.). Randolph: Author.
- National Autism Center. (2015b). *Findings and conclusions: National standards project, phase 2*. Randolph: Author.
- Parker, D., & Kamps, D. (2011). Effects of task analysis and self-monitoring for children with autism in multiple social settings. *Focus on Autism and Other Developmental Disorders*, *26*(3), 131–142.
- Pierson, M. R., & Glaeser, B. C. (2005). Extension of research on social skills training using comic strip conversations to students without autism. *Education and Training in Developmental Disabilities*, *40*(3), 279–284.
- Pierson, M. R., & Glaeser, B. C. (2007). Using comic strip conversations to increase social satisfaction and decrease loneliness in students with autism spectrum disorder. *Education and Training in Developmental Disabilities*, *42*(4), 460–466.
- Reynhout, G., & Carter, M. (2011). Evaluation of the efficacy of social stories™ using three single subject metrics. *Research in Autism Spectrum Disorders*, *36*, 445–469.
- Rhodes, C. (2014). Do social stories help to decrease disruptive behaviour in children with autistic spectrum disorders? A review of the published literature. *Journal of Intellectual Disabilities*, *18*, 35–50.
- Roth, M. E., Gillis, J. M., & DiGennaro Reed, F. D. (2014). A meta-analysis of behavioral interventions for adolescents and adults with autism spectrum disorders. *Journal of Behavioral Education*, *23*(2), 258–286.
- Styles, A. (2011). Social stories™: Does the research evidence support the popularity? *Educational Psychology in Practice*, *27*, 415–436.
- Van Houten, R., Axelrod, S., Bailey, J. S., Favell, J. E., Foxx, R. M., Iwata, B. A., & Lovaas, O. I. (1988). The right to effective behavioral treatment. *Journal of Applied Behavior Analysis*, *21*(4), 381–384.
- Wong, C., Odom, S. L., Hume, K., Cox, A. W., Fetting, A., Kucharczyk, S., et al. (2013). *Evidence-based practices for children, youth, and young adults with autism Spectrum disorder*. Chapel Hill: The University of North Carolina, Frank Porter Graham Child Development Institute, Autism Evidence-Based Practice Review Group.
- Wong, C., Odom, S. L., Hume, K. A., Cox, A. W., Fetting, A., Kucharczyk, S., Brock, M. E., Plavnick, J. B., Fleury, V. P., & Schultz, T. R. (2015). Evidence-based practices for children, youth, and young adults with autism spectrum disorder: A comprehensive review. *Journal of Autism and Developmental Disorders*, *45*(7), 1951–1966. <https://doi.org/10.1007/s10803-014-2351-z>.
- Zimmerman, K. N., & Ledford, J. R. (2017). Beyond ASD: Evidence for the effectiveness of social narratives. *Journal of Early Intervention*, *39*(3), 199–217. <https://doi.org/10.1177/105381511770900>.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Journal of Developmental & Physical Disabilities is a copyright of Springer, 2020. All Rights Reserved.