

Systematic review of collateral effects of focused interventions for children with autism spectrum disorder

Autism & Developmental Language
Impairments
Volume 2: 1–22
© The Author(s) 2017
DOI: 10.1177/2396941517737536
journals.sagepub.com/home/dli



Katherine Ledbetter-Cho

University of Texas at Austin, USA; Texas State University, USA

Russell Lang

Texas State University, USA

Laci Watkins

University of Alabama, USA

Mark O'Reilly and Claudia Zamora

University of Texas at Austin, USA

Abstract

A collateral intervention effect refers to changes in behaviors which were not directly targeted during intervention. Using predetermined search and inclusion procedures, this systematic review identified 46 studies involving children with autism spectrum disorder and 14 desirable collateral effects across multiple domains of functioning. Collateral effects were associated with: (a) interventions involving naturalistic behavioral strategies; (b) participants with limited communication and/or cognitive deficits; (c) performance deficits (i.e. there was some evidence of the collateral behavior in baseline); and (d) interventions directly targeting play, communication, joint attention, and/or stereotypy. Overall, this systematic review indicates that collateral effects arising from focused interventions warrant consideration by practitioners during intervention planning and require additional research to identify mechanisms responsible for the observed changes.

Keywords

Collateral, untargeted, response generalization, novel, autism

The diagnostic criteria for autism spectrum disorder (ASD) consist of impairment in social communication and restricted interests and patterns of behavior (American Psychiatric Association [APA], 2013). Although not part of the ASD diagnostic criteria, challenging behavior (e.g. aggression, self-injury) and co-morbid diagnoses (e.g. anxiety disorder, intellectual disability) are more prevalent in samples of children with ASD (Mannion & Leader, 2013; Matson & Nebel-Schwalm, 2007). These characteristics often have deleterious effects across a variety of domains (e.g. language, play, daily-living skills) which, in the absence of intervention, present obstacles to forming

social relationships, educational attainment, employment, and autonomy throughout life (Chamak & Bonniau, 2016; Henninger & Taylor, 2012). Given the pervasiveness of skill deficits and behavioral excesses that may warrant intervention in children with ASD, interventions that occasion concomitant improvements in behaviors not directly targeted during intervention (collateral effects) may offer desirable intervention efficiency (Koegel, Koegel, & McNERney, 2001; McConnell, 2002; Pauwels, Ahearn, & Cohen, 2015).

Currently, intervention approaches most commonly associated with improvements across skill domains for children with ASD tend to be intensive (e.g. 20 to 40

Corresponding author:

Katherine Ledbetter-Cho, Department of Special Education, University of Texas at Austin, 1912 Speedway Stop, D5700 Austin, TX 78712, USA.
Email: kledbetter-cho@utexas.edu



hours per week), initiated early in life, and involve multiple intervention components that directly target a comprehensive set of behaviors (e.g. Lang, Hancock, & Singh, 2016; Virués-Ortega, 2010). Comprehensive and intensive intervention has been demonstrated to improve areas directly related to ASD diagnostic criteria (e.g. social communication), ameliorate common comorbidities (e.g. challenging behavior), and may even result in more typical neurological functioning (e.g. Dawson et al., 2012; Reichow, Barton, Boyd, & Hume, 2014; Ryberg, 2015; Vismara & Rogers, 2010). Unfortunately, many children with ASD and their families are confronted with a lack of available service providers with expertise in comprehensive intervention, prohibitive intervention costs, and comorbid health conditions that interrupt or preclude intensive intervention (Jacobson & Mulick, 2000; Thomas, Ellis, McLaurin, Daniels, & Morrissey, 2007; Vohra, Madhavan, Sambamoorthi, & Peter, 2014). These factors necessitate consideration of other less comprehensive intervention options (Pickard & Ingersoll, 2016).

As opposed to targeting a broad range of behaviors across multiple domains via comprehensive intervention, another option is a focused approach to intervention. Focused interventions involve the selection of a specific target behavior (e.g. initiating play with a peer) or a small set of related target behaviors (e.g. social initiations and responses) and then the development of an intervention that focuses on the selected behaviors (O'Reilly, Falcomata, Kang, & Fragale, 2014). Selection of target behaviors and focused intervention components is based on a multitude of considerations including developmental appropriateness, ecological validity, assessment of the child's existing skills and preferences, and family input regarding treatment priorities and preferences (Baer, Wolf, & Risley, 1987; Lifter, Ellis, Cannon, & Anderson, 2005; Lifter, Sulzer-Aaaroff, Anderson, & Cowdery, 1993). An additional consideration which may increase the efficacy of focused interventions is the selection of behaviors and intervention procedures that have been shown to produce collateral, or untargeted, skill improvements (McConnell, 2002). A number of terms used in the ASD intervention literature are related to collateral effects including response generalization, behavioral cusp, and pivotal response.

Response generalization refers to a generalized behavior change wherein a change in a targeted behavior results in a change in a nontargeted behavior that shares the same operant function, similar discriminative stimuli, or related topographies (Cooper, Heron, & Heward, 2007; Kazdin, 1994; Stewart, McElwee, & Ming, 2013). For example, an intervention designed to teach a specific form of play behavior that results in the child acquiring the target play skill as well as a

play skill not directly targeted during intervention could be described as having resulted in response generalization (e.g. Lang et al., 2014).

The term behavioral cusp references a wider spread of collateral effects than is typically associated with the term response generalization. The concept of behavioral cusp refers to cases where change in a target behavior profoundly influences many nontargeted behaviors across multiple domains (Smith, McDougall, & Edelen-Smith, 2006). Bosch and Fuqua (2001) described behavioral cusps in the context of interventions that result in: (a) acquisition of a target behavior that enables exposure to new contingencies of reinforcement and novel environments; (b) socially valid behavior change; (c) increased ability to originate, produce, or create (i.e. generativeness); and (d) displacement of inappropriate behavior. Interventions targeting joint attention that also demonstrate collateral improvements in language which in turn lead to a reduction in challenging behavior are putative examples of this concept (White et al., 2011).

Pivotal responses are a group of specific intervention targets with the potential to occasion a broad range of concomitant behavior changes similar to a behavioral cusp. Acquisition of a pivotal response is theorized to reduce learned helplessness and increase motivation to respond to social and instructional stimuli (Koegel, Ashbaugh, & Koegel, 2016). Target behaviors considered pivotal responses include social initiations, attending to multiple features of a stimulus, and self-management skills (e.g. Koegel & Koegel, 2006; Koegel & Wilhelm, 1973; Koegel & Mentis, 1985; Schreibman & Stahmer, 2014). For example, a child who is taught to initiate a social interaction with peers may acquire other social skills and novel play behaviors as a result of increased peer interaction. The extent to which collateral effects associated with the acquisition of pivotal responses are a product of the specific pivotal responses targeted, the intervention components utilized, or a combination is not yet clear (Cadogan & McCrimmon, 2015).

Several previous reviews have addressed collateral effects but only in the context of a specific intervention package or a specific target behavior. For example, Verschuur, Didden, Lang, Sigafoos, and Huskens (2014) reviewed 43 studies investigating pivotal response treatment (PRT) and reported that targeted increases in social initiations were associated with collateral improvements in language, play skills, and challenging behavior. Similarly, a meta-analysis of the Picture Exchange Communication System (PECS) reported collateral improvements in spoken language, socialization, and challenging behavior (Ganz, Davis, Lund, Goodwyn, & Simpson, 2012). With regard to focusing on a specific target behavior, Lanovaz,

Robertson, Soerono, and Watkins (2013) reviewed 60 studies targeting a reduction in stereotypy and reported that a decrease in the targeted form of stereotypy occasioned a desirable increase in adaptive behavior in most cases and an undesirable increase in other forms of stereotypy in a few cases. Finally, White et al. (2011) reviewed 27 studies that measured joint attention. When intervention targeting joint attention was effective, collateral improvements in social initiations, imitation, play, and speech were often reported.

This systematic review extends previous reviews by focusing on collateral intervention effects without restricting included studies to a specific focused intervention package or class of target behaviors. Given the importance of beginning intervention early in life (Pickles et al., 2016), the current systematic review aims to identify collateral effects demonstrated in early childhood intervention studies. Further, with the exception of the meta-analysis of PECS (Ganz et al., 2012), effect size estimates for collateral changes have not been calculated in previous reviews. The goals of the present systematic review are to: (a) identify collateral effects that have been reported in intervention research involving children with ASD; (b) evaluate the methodological rigor and calculate effect size estimates of the targeted and collateral behavior changes; (c) identify characteristics of participants, interventions, and target behaviors that may influence collateral effects; and (d) discuss implications for practice and research.

Methods

Protocol registration and PRISMA guidelines

The protocol for this systematic review was registered with the PROSPERO International prospective register of systematic reviews and was prepared in accordance with PRISMA guidelines (Ledbetter-Cho, Lang, Watkins, & O'Reilly, 2015; Moher, Liberati, Tetzlaff, & Altman, 2009).

Search strategy

A systematic search of four electronic databases was conducted, including Educational Resources Information Center (ERIC), Medline, Psychology and Behavioral Sciences Collection, and PsychINFO. Searches consisted of combinations of terms referring to collateral intervention effects (i.e. non-targeted or nontargeted or untargeted or unanticipated or collateral or concomitant or behavioral cusp or pivotal response or ancillary or response generalization); terms related to diagnosis (i.e. autism* or ASD or Asperger* or pervasive developmental disorder*); and

terms suggesting an intervention study (i.e. intervention or treatment or program or train*). Because terms related to collateral behaviors may not appear in an article's title, abstract or keyword list, we set search parameters to "open field." An open field search identifies articles containing the search term anywhere in the text (not limited to title, abstract, or key terms). Publication date was also unrestricted but studies were limited to those written in English and published in peer-reviewed journals. This database search procedure yielded 710 studies. Next, secondary searches of included articles and of previous literature reviews were conducted. Finally, hand searches of journals that often publish intervention research with children with ASD (e.g. *Journal of Autism and Developmental Disorders*) were conducted. The first and second author initially applied the inclusion criteria to the corpus of studies resulting from the search procedures. The third author then independently screened articles identified for inclusion and interrater agreement reached 98%. Based on recommendations from the Cochrane Collaboration, the disagreement was resolved by discussion among the authors (Higgins & Green, 2011). Figure 1 depicts the search and screening process.

Study selection

An intervention study was required to meet predetermined criteria to be included. First, the intervention

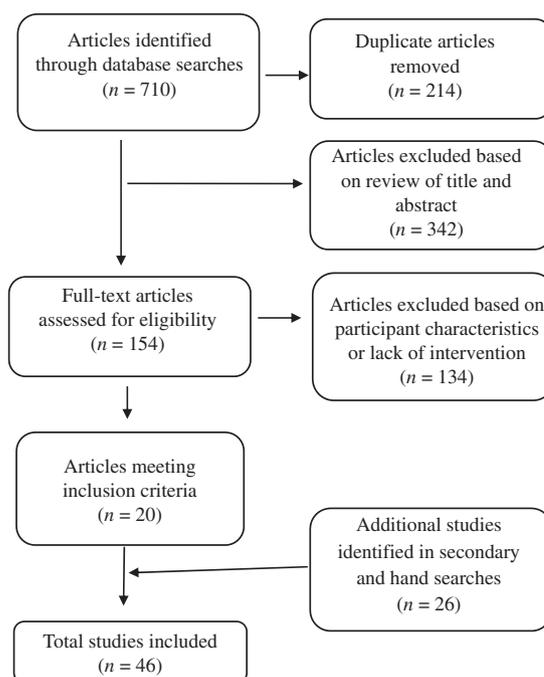


Figure 1. Flowchart of included studies.

had to be delivered to at least one child (birth through 8 years old) diagnosed with ASD, Autistic Disorder, Asperger's Syndrome, or Pervasive Developmental Disorder-Not Otherwise Specified (PDD-NOS). If a study included some participants that met criteria and others that did not, only data pertaining to participants meeting the criteria were considered (e.g. Lanovaz et al., 2014; Lee & Odom, 1996). The study was excluded if data from participants meeting criteria could not be disaggregated from other participants' data (e.g. Karaaslan, Diken, & Mahoney, 2013). Second, studies involving comprehensive early intensive intervention packages (e.g. the Early Start Denver Model) and those involving biomedical or physiological procedures (e.g. exercise, dietary manipulations and chelation) were excluded because it was not possible to disaggregate collateral effects from the targeted behavior changes (e.g. Celiberti, Bobo, Kelly, Harris, & Handleman, 1997; Lovaas, Koegel, Simmons, & Long, 1973; Rogers et al., 2012). For example, studies involving sensory integration therapy were excluded because the purported mechanism of action involves changes in sensory processing via neuroplasticity and such a change (if it occurs) would be expected to have a wide spread of effects (e.g. Reichow, Barton, Sewell, Good, & Wolery, 2010). Third, a study had to clearly identify at least one target behavior (e.g. a specific communication, play, or social skill) measured by direct observation and describe intervention procedures focused directly on that target behavior (c.f., McEvoy et al., 1988). Fourth, the study had to include data indicating a change in a behavior that was not directly targeted by an intervention component or procedure (i.e. collateral effect). For example, studies involving Functional Communication Training (FCT) that reported a decrease in challenging behavior and an increase in an alternative targeted communication behavior would be excluded because the alternative communication behavior is prompted and differentially reinforced and challenging behavior is put on extinction: as such, there is an intervention component directly aimed at both dependent variables (Carr & Durrand, 1985). Similarly, if engagement in the target behavior displaced another behavior because the two were physically incompatible (e.g. the study's operational definition of on-task behavior required the child to stop engaging in stereotypy), the study was excluded.

Finally, included studies had to utilize an experimental group design or demonstrate experimental control for at least one target or collateral behavior in a single-case design (SCD). In some cases, experimental control in a SCD was demonstrated with either the target behavior or the collateral behavior but not both. For example, some studies targeted a specific behavior for

improvement (e.g. communication) but only reported data on the collateral variable (e.g. social engagement; Koegel, Vernon, & Koegel, 2009). If experimental control was demonstrated for the collateral variable, the study was included. Other SCD studies demonstrated experimental control over the target behavior but reported collateral effects as averages across phases or participants, precluding the visual analysis of trend and variability necessary to evidence experimental control (e.g. Goldstein & Cisar, 1992). In those cases, at a minimum, the study had to measure the collateral behavior pre- and post-intervention in addition to demonstrating experimental control with the target behavior to be included. Differences in experimental control for target and collateral behaviors were accounted for when coding research rigor (see Data Extraction and Coding). If experimental control was compromised for both targeted and collateral effects by the exclusion of participants older than 8 years or without ASD (e.g. exclusion of a participant in a multiple baseline across participants design), the study was excluded (e.g. Thorp, Stahmer, & Schreibman, 1995). These exclusions ensure a minimum degree of rigor among included studies.

Data extraction and coding

Table 1 displays data extracted from the 46 studies including: (a) number, gender, age, and functioning level of participants; (b) intervention procedures, setting, practitioner, and dosage (e.g. total hours of intervention); and (c) effect size estimates and research rigor classification for each target and collateral behavior. When two studies reported intervention outcomes for the same group of participants, data for both studies were consolidated and reported as a single entry in the table (e.g. Wichnick, Vener, Keating, & Poulson, 2010; Wichnick, Vener, Pyrtex, & Poulson, 2010). If a study contained more than one experiment, only the experiments meeting inclusion criteria were included (e.g. Nuzzolo-Gomez, Leonard, Oritz, Rivera, & Greer, 2002).

The first author developed a coding manual for data collection and analysis. After data collection was complete, the third author independently verified 30% of included studies. Agreements were defined as a match between the two coders, with effect size estimates required to match to the hundredths place in order to be scored as an agreement. Interrater agreement was 95.6% and was calculated by dividing the number of agreements by the total number of agreements plus disagreements and multiplying by 100.

Participant functioning level was categorized as lower, medium, or higher functioning according to the framework provided by Reichow and Volkmer

Table 1. Summary of included studies.

Citation	N Participants (n female); Age Range ^a ; Functioning Level	Intervention	Intervention Setting; Practitioner; Dosage	Target Behavior Change: Effect Size Estimates (PND, IRD, and NAP) or Cohen's d; Rigor Classification	Collateral Behavior Change: Effect Size Estimates (PND, IRD, and NAP) or Cohen's d; Rigor Classification
Studies targeting communication and/or social skills (N = 31)					
Baker (2000)	N = 3 (1 f) Age 5:5 to 6:10 Medium	Perseverative interests incorporated into games	Clinic; Researcher; 30– 40 min × up to 14 sessions (up to 9.3 hrs.)	Social play: (.97, .96, .99) Rigor: Adequate	Joint attention: (.91, .96, .98) Rigor: Adequate Stereotypy: (NR, .44, .69) Rigor: Weak
Charlop-Christy, Carpenter, Le, LeBlanc, and Kellert (2002)	N = 2 Age 3:8 to 5:9 Lower	PECS	Clinic; Therapist; 15 min 2 × per wk. (up to 2.9 hrs.)	Requests with picture symbols: (NR) Rigor: Weak	Tantrums: (.25, .54, .76) Disruptions: (.10, .44, .71) Rigor (all): Weak
Charlop and Trasowech (1991)	N = 3 Age 7:9 to 8:7 Medium and higher	Prompting and reinforcement	Home and clinic; Parents and therapist; Daily incidental trials, 4–16 wks.	Social language: (.54, .78, .91) Rigor: Weak	Novel/varied speech responses: (NR) Rigor: Weak
Delfs et al. (2014)	N = 4 Age 3–8 Lower and medium	Discrete trial teaching	Clinic-based classroom; Researcher; 2–5 × per wk. (up to 58 sessions)	Receptive identification: (.80, .83, .95) Expressive identification: (.90, .90, .96) Rigor (all): Adequate	Generalization to receptive: (.88, .92, .97) Generalization to expressive: (.65, .70, .85) Rigor (all): Adequate
Gianoumis, Seiverling, and Sturmei (2012)	N = 6 (3 f) Age 3–4 Lower and medium	Naturalistic behavioral strategies	Classroom; Teacher; Up to 5 sessions	Vocal utterances: (.55, .75, .79) Rigor: Strong	Challenging behavior: (.55, .75, .85) Rigor: Adequate
Goldstein and Cisar (1992)	N = 2 Preschool age Lower	Script training	Empty room; Teacher; 15 min × 22 sessions (5.5 hrs.)	Scripted play verbalizations: (1) Rigor: Adequate	Untaught/novel play verbalizations: (NR) Untaught/novel play actions: (NR) Rigor (all): Weak
Ingersoll and Schreibman (2006)	N = 5 (2 f) Age 2:4 to 3:9 Lower and medium	Naturalistic behavioral strategies	Empty room; Researcher; 20 min × up to 29 sessions (up to 9.6 hrs.)	Total object imitation: (.74, .75, .92) Rigor: Adequate	Imitative vocal utterances: (.45, .40, .75) Spontaneous vocal utterances: (.30, .43, .67) Total play: (.23, .47, .75) Spontaneous play: (.15, .35, .68) Coordinated joint attention: (.41, .53, .79) Rigor (all): Weak
Kasari, Freeman, and Paparella (2006) and	Joint attention group: N = 20 (5 f); age 3–4 Lower and medium	Naturalistic behavioral strategies	Clinic; Researcher; 30 min 5 × per wk., 5–6 wks. (up to 15 hrs.)	Joint attention group versus con- trol group Coordinated joint attention	JA group versus control group Vocal utterances (.59) in Cohen's d SP group versus control group

(continued)

Table 1. (continued)

Citation	N Participants (n female); Age Range ^a ; Functioning Level	Intervention	Intervention Setting; Practitioner; Dosage	Target Behavior Change: Effect Size Estimates (PND, IRD, and NAP) or Cohen's <i>d</i> ; Rigor Classification	Collateral Behavior Change: Effect Size Estimates (PND, IRD, and NAP) or Cohen's <i>d</i> ; Rigor Classification
Kasari, Paparella, Freeman, and Jahromi (2008)	N Participants (n female); Age Range ^a ; Functioning Level <i>Symbolic play group</i> : N = 21 (5 f); age 3–4 Lower and medium <i>Control group</i> : N = 17 (2 f); age 3–4 Lower and medium			looks (.50), gives (.36), shows (.52), child-initiated joint attention (.79), and joint attention responses (1.13) in Cohen's <i>d</i> <i>SP group versus control group</i> Play level in assessment (.45) and with parent (.74), and symbolic play types (.50) in Cohen's <i>d</i>	Coordinated joint attention looks (.66) and vocal utterances (.71) in Cohen's <i>d</i> Rigor (<i>all</i>): Strong
Koegel et al. (1998)	N = 3 (1 f) Age 3:8 to 5:5 Lower and medium	Naturalistic behavioral strategies	Clinic; Researcher; 30 min × up to 16 sessions (up to 8 hrs.)	Rigor (<i>all</i>): Strong Social question-asking: (.97, .97, .98)	Expressive labeling of stimulus items: (1) Rigor: Adequate
Koegel, Bradshaw, Ashbaugh, and Koegel (2014)	N = 3 Age 3:2 to 3:7 Medium	Naturalistic behavioral strategies	Clinic; Researcher; 10 hrs. per week × 10 mths. (up to 400 hrs.)	Rigor: Adequate Social question-asking: (.87, .86, .95)	Untaught questions: (.73, .73, .86) Rigor: Adequate
Koegel et al. (1992)	N = 3 (1 f) Age 3:4 to 4:6 Lower	Naturalistic behavioral strategies	Clinic; Researcher; 10 min × up to 10 sessions (up to 1.6 hrs.)	Rigor: Adequate Vocal utterances: (NR) Rigor: Weak	Disruptive behavior: (1) Rigor: Strong
Koegel et al. (2009)	N = 3 Age 3:2 to 3:5 Lower and medium	Naturalistic behavioral strategies with social reinforcers	Home; Researcher or parent; 2 hrs. × up to 7 sessions per condition (up to 14 hrs. per condition)	Vocal utterances: (NR) Rigor: Weak	Social engagement: (1) Orienting: (1) Rigor (<i>all</i>): Strong
Lang et al. (2014)	N = 3 (1 f) Age 3:6 to 3:10 Lower and medium	Prompting and reinforce- ment with lag schedule for two	Special education classroom; Teacher; 5 min × up to 16 sessions (up to 1.3 hrs.)	Functional play: (1) Rigor: Strong	Stereotypy: (.92, .88, .99) Rigor: Adequate
Ledbetter-Cho et al. (2015)	N = 3 Age 4:9 to 6:3 Medium and higher	Script training	Clinic; Researcher; 7 min × up to 32 sessions (up to 3.7 hrs.)	Scripted initiations: (NR) Rigor: Weak	Responses: (.41, .37, .70) Rigor: Weak Unscripted initiations: (.73, .64, .87) Novel initiations: (.36, .53, .81) Novel responses: (.33, .37, .70) Novel content changes: (.34, .33, .66) Rigor (<i>rest</i>): Adequate
Lee and Odom (1996)	N = 1 Age 7 Lower	Peer-mediated social interaction	Classroom; Peers; 6– 10 min × 20 sessions (up to 3.3 hrs.)	Social interaction: (1) Rigor: Adequate	Stereotypy: (.85, .90, .97) Rigor: Weak

(continued)

Table 1. (continued)

Citation	N Participants (n female); Age Range ^a ; Functioning Level	Intervention	Intervention Setting; Practitioner; Dosage	Target Behavior Change: Effect Size Estimates (PND, IRD, and NAP) or Cohen's <i>d</i> ; Rigor Classification	Collateral Behavior Change: Effect Size Estimates (PND, IRD, and NAP) or Cohen's <i>d</i> ; Rigor Classification
MacDonald, Sacramone, Mansfield, Wiltz, and Ahearn (2009)	N = 2 Age 5-7 Higher	Video modeling	Empty room; Researcher; ~5 min up to 22 sessions (up to ~1.8 hrs.)	Social play actions: (.95, .95, .96) Social play verbalizations: (.95, .95, .98) Rigor (all): Adequate	Untaught play verbalizations: (NR) Rigor: Weak
Nuzzolo-Gomez et al. (2002); Exp. 2	N = 3 (1 f) Age 4-7 Lower and medium	Prompting and reinforcement	Self-contained classroom; Researcher; Up to 18 sessions	Functional play: (.42, .67, .81) Rigor: Weak	Stereotypy: (.25, .67, .83) Rigor: Adequate
Plavnick and Ferreri (2012)	N = 3 (1 f) Age 4:6 to 6:6 Lower	Mand training (establish- ing operation contrived)	Empty room; Researcher; 15 min × up to 15 sessions per condition (up to 3.75 hrs. per condition)	Requests: (NR) Rigor: Weak	Orienting: (.85, .61, .80) Compliance: (.66, .48, .72) Challenging behavior: (.85, .69, .91) Rigor (all): Adequate
Pollard, Betz, and Higbee (2012)	N = 3 (1 f) Age 4-7 Higher	Script training and mul- tiple exemplars	Empty room; Researcher; 2- 5 min × up to 48 sessions (up to 4 hrs.)	Scripted bids for joint atten- tion: (NR) Rigor: Weak	Unscripted bids for joint attention: (.73, .91, .95) Rigor: Adequate
Rocha et al. (2007)	N = 3 (1 f) Age 2:2 to 3:6 Lower and medium	Naturalistic behavioral strategies	Clinic; Parent; 20 min × up to 50 sessions (up to 19.3 hrs.)	Joint attention responses: (.54, .81, .95) Rigor: Adequate	Joint attention initiations: (NR) Rigor: Weak
Sarokoff et al. (2001)	N = 1 Age 8 Medium	Script training	Clinic; Researcher; 3 min × 81 sessions (4.05 hrs.)	Scripted initiations: (1) Rigor: Adequate	Unscripted statements: (.49, .60, .80) Rigor: Adequate
Vernon, Koegel, Dauterman, and Stolen (2012)	N = 3 Age 2:4 to 4:3 Lower	Naturalistic behavioral strategies with social reinforcers	Home; Parent; 10 min × 16 sessions (2.6 hrs.)	Vocal utterances: (.79, .86, .94) Rigor: Adequate	Eye contact: (.89, .89, .96) Rigor: Adequate
Vismara and Lyons (2007)	N = 3 Age 2:2 to 3:2 Lower	Perseverative interests with naturalistic behavioral strategies	Clinic or home; Parent; 2.5 hr. sessions (up to 30 hrs.)	Vocal utterances: (NR) Rigor: Weak	Joint attention initiations: (.48, .58, .72) Rigor: Adequate
Whalen, Liden et al. (2006) Study 2	N = 4 Age 3:4 to 4:3 Lower and medium	Computer software	Home; Parent; 15 min 3 × per wk., 8 wks. (6 hrs.)	Language targets on computer: (NR) Rigor: Weak	Spontaneous vocal utterances: (.67, .66, .86) Looking to adults: (.75, .75, .90) Inappropriate behavior: (.91, .91, .95) Inappropriate language: (.83, .91, .83) Rigor (all): Weak
Whalen, Schreibman, and Ingersoll (2006)	N = 4 Age 4 to 4:4 Medium	Naturalistic behavioral strategies	Clinic; Researcher; 25 min 9 × per wk., ~10 wks. (~37.5 hrs.)	Joint attention responses: (.61, .69, .88) Joint attention initiations: (.91, .91, .95) Rigor (all): Adequate	Imitation: (NR) Vocal utterances: (NR) Rigor (all): Weak

(continued)

Table 1. (continued)

Citation	N Participants (n female); Age Range ^a ; Functioning Level	Intervention	Intervention Setting; Practitioner; Dosage	Target Behavior Change: Effect Size Estimates (PND, IRD, and NAP) or Cohen's <i>d</i> ; Rigor Classification	Collateral Behavior Change: Effect Size Estimates (PND, IRD, and NAP) or Cohen's <i>d</i> ; Rigor Classification
Wichnick, Vener, Keating, et al. (2010) and Wichnick, Vener, Pyrtek, et al. (2010)	N = 3 (1 f) Age 4–6 Medium	Script training	Classroom; Researcher; Up to 178 sessions for initiations and 80 sessions for responses	Scripted initiations: (.91, .90, .95) Scripted responses: (.88, .89, .94) Rigor (all): Strong	Novel initiations: (.12, .22, .61) Novel responses: (.20, .41, .70) Rigor (all): Adequate
Wynn and Smith (2003)	N = 6 Age 3:11 to 6:4 Lower, medium and higher	Discrete trial teaching	Home; Researcher; Up to 60 sessions	Receptive identification: (.90, .90, .96) Expressive identification: (.93, .92, .96) Rigor (all): Adequate	Generalization to receptive: (.75, .75, .81); Generalization to expressive: (.51, .51, .66) Rigor (all): Adequate
Yoder and Stone (2006a, 2006b)	Responsivity Prelinguistic Milieu group: N = 17 (3 f); age 2– 5 Lower PECS group: N = 19 (3 f); age 2–5 Lower	Responsive Education and Prelinguistic Milieu Teaching or PECS	Clinic; Researcher; 20 min × 72 sessions (24 hrs.)	RPMT group versus PECS group: Object exchange turns/pre- linguistic joint attention acts (.97) in Cohen's <i>d</i> PECS group versus RPMT group: Requests with picture sym- bols (NR) Rigor (all): Strong	PECS group versus RPMT Nonimitative spoken acts (.63), dif- ferent nonimitative spoken acts: (.50) in Cohen's <i>d</i> ; Joint attention initiations for partici- pants who entered with one or fewer initiations (NR) Rigor (all): Strong
Studies targeting restricted, repetitive behaviors and/or interests (N = 8)					
Ahearn, Clark, MacDonald, and Chung (2007)	N = 3 (2 f) Age 3–7 Lower and medium	Vocal RIRD	Empty room; Therapist/ tea- cher; ~5 min, up to 11 ses- sions (up to ~1 hr.)	Vocal stereotypy: (.92, .96, .98) Rigor: Strong	Vocal utterances: (.70, .60, .63) Rigor: Adequate
Ahrens et al. (2011)	Experiment 1: N = 2 Age 4–6 Lower and higher Experiment 2: N = 2 Age 4–5 Medium	Motor RIRD or vocal RIRD	Empty room, home, or clinic; Therapist ~5 min × up to 54 sessions (up to ~4.5 hrs.) Clinic; Therapist ~5 min × up to 45 sessions (up to ~3.75 hrs.)	Vocal stereotypy: (.22, .69, .86); Rigor: Adequate Vocal or motor stereotypy: (.86, .85, .96) Rigor: Adequate	Vocal utterances: (.44, .65, .84) Rigor: Adequate Vocal utterances: (.17, .25, .33) Rigor: Weak
Cook, Rapp, Gomes, Frazer, and Lindblad (2014)	N = 1 Age 5 Lower	Verbal reprimands/stimu- lus control	Clinic; Researcher; 5 min × 21 sessions (1.75 hrs.)	Motor stereotypy: (1) Rigor: Weak	Undesirable increase in untargeted stereotypy: (.90, .75, .87) Requests: (.61, .55, .79) Rigor (all): Weak
Greer and Han (2015)	N = 3 Age 5:3 to 5:7 Lower and medium	Conditioning new reinforcers (via stimu- lus pairing) Punishment	Self-contained classroom; Researcher; Up to 13 sessions	Observing conditioned reinfor- cers: (NR) Rigor: Weak Stereotypy: (.78, .92, .98) Rigor: Adequate	Identical and non-identical matching: (1) Rigor: Adequate Functional play: (.17, .73, .93) Rigor: Adequate

(continued)

Table 1. (continued)

Citation	N Participants (n female); Age Range ^a ; Functioning Level	Intervention	Intervention Setting; Practitioner; Dosage	Target Behavior Change: Effect Size Estimates (PND, IRD, and NAP) or Cohen's <i>d</i> ; Rigor Classification	Collateral Behavior Change: Effect Size Estimates (PND, IRD, and NAP) or Cohen's <i>d</i> ; Rigor Classification
Koegel, Firestone, Kramme, and Dunlap (1974)	N = 2 (1 f), Age 6–8 Lower		Clinic; Researcher; 5 min × up to 17 sessions (up to 1.4 hrs.)		
Lang et al. (2009)	N = 1 f Age 8 Lower	AOC prior to play session	Empty room; Researcher; 15 min × 6 sessions (1.5 hrs.)	Stereotypy: (.83, .66, .81) Rigor: Adequate	Challenging behavior: (.66, .66, .80) Functional play: (1, .5, .75) Rigor (<i>dif</i>): Adequate
Lang et al. (2010)	N = 4 (2 f) Age 4–7 years Lower	AOC prior to play session	Empty room; Teacher; ~34 min × up to 6 sessions (up to ~3.4 hrs.)	Stereotypy: (.90, .68, .86) Rigor: Strong	Challenging behavior: (.81, .72, .86) Functional play: (.68, .41, .63) Rigor (<i>dif</i>): Adequate
Lanovaz et al. (2014)	N = 1 Age 4 NR	Noncontingent music	Setting NR; Researcher; 10 min, 6 sessions (1 hr.)	Vocal stereotypy: (1) Rigor: Adequate	On-task behavior: (.66, .33, .66) Rigor: Weak
Studies targeting other skills (N = 7)					
Bennett, Reichow, and Wolery (2011); Study 2	N = 1 f Age 3:10 Lower	Visual work system	Classroom; Researcher; 3 min × 20 sessions (1 hr.)	On-task in academics: (.90, .93, .99) Rigor: Adequate	Stereotypy: (NR, .71, .87) Rigor: Weak
Davis et al. (2009)	N = 1 Age 4 Medium	Abolishing operation component	Empty room; Researcher; 20 min × 8 sessions (2.6 hrs.)	Challenging behavior: (.87, .75, .89) Rigor: Adequate	Requests: (.75, .50, .77) Rigor: Adequate
Kamps, Barbetta, Leonard, and Delquadri (1994)	N = 2 Age 8 Medium and higher	Peer-mediated instruction	Classroom; Peers; 25– 30 min × up to 40 sessions plus 2.25 hr. initial training (up to 22.25 hrs.)	Reading comprehension: (.39, .54, .87) Rigor: Weak	Social interaction: (.53, .59, .89) Rigor: Weak
Kinney, Vedora, and Stromer (2003); Phase 4	N = 1 f Age 8 Medium	Matrix training	Home/school; Teacher/ Researcher; 18 sessions	Spelling words: (NR) Rigor: Weak	Untaught spelling words: (1) Rigor: Adequate
Krantz, MacDuff, and McClannahan (1993)	N = 3 Age 6–8 Lower	Visual work system	Home; Parent; 17–22 sessions (up to 90 hrs.)	Daily-living skills: (1) Rigor: Strong	Disruptive behavior: (1) Rigor: Strong
Pierce and Schreibman (1994)	N = 2 Age 6–8 Lower	Visual work system	Clinic/ home; Therapist; Up to 18 sessions	Daily-living skills: (1) Rigor: Strong	Inappropriate behavior: (.88, .88, .95) Rigor: Adequate
Soutor et al. (1994)	N = 1 Age 3:8 Lower	Reinforcement	Self-contained classroom; Researcher; 20 min × up to 26 sessions (up to 8.6 hrs.)	Compliance: (.80, .84, .96) Rigor: Adequate	Attending: (.84, .84, .93) Rigor: Adequate Vocal utterances: (.11, .11, .51) Rigor: Weak

^aYears:months; f: female; AOC: abolishing operation component; Exp.: experiment; IRD: improvement rate difference; NAP: nonoverlap of all pairs; NR: not reported or not ratable; PECS: picture exchange communication system; PND: percent nonoverlapping data; RIRD: response interruption and redirection.

(2010). Lower functioning refers to participants with very limited vocal communication skills or an IQ below 55. Participants classified as medium functioning had emerging vocal communication or an IQ between 55 and 85. Those classified as higher functioning displayed age-appropriate vocal communication and an average or above-average IQ.

Research rigor was coded according to criteria outlined by Reichow, Volkmar, and Cicchetti (2008) which has precedence in reviews of ASD intervention research (e.g. Siegel & Beaulieu, 2012; Whalon, Conroy, Martinez, & Werch, 2015). Specifically, methodological strength was coded as strong, adequate, or weak dependent upon the number of primary and secondary quality indicators met. Primary quality indicators include clear descriptions of participant characteristics, operational definitions of independent and dependent variables, and demonstration of experimental control in SCD or appropriate statistical analyses and power in group designs. Secondary quality indicators include adequate interobserver agreement (IOA), blind raters, treatment fidelity, generalization, maintenance, and social validity. Dependent variables classified as having strong methodological rigor met all primary quality indicators and at least three secondary quality indicators in SCD or four in group designs. Adequate rigor was assigned to variables that evidenced at least four primary quality indicators and two secondary quality indicators. Variables considered to have weak rigor met fewer than four primary quality indicators or fewer than two secondary quality indicators.

Several of the quality indicators described above focus on the rigor of data collection and analysis procedures. Because those procedures sometimes differed for target behaviors and collateral behaviors within the same study (e.g. sufficient baseline data collected for target behavior but not collateral behavior), multiple quality indicator scores were calculated per study. Scores were calculated first by applying Reichow et al.'s (2008) quality indicators to the target dependent variable. However, given the differences in procedures used for target and collateral behaviors, a rigor classification based on target behaviors should not be conflated with the certainty of evidence for the collateral behavior variables. Therefore, separate rigor scores were also calculated for collateral behavior dependent variables by applying the same Reichow et al. criteria to those data. This approach empowers a more nuanced consideration of the certainty of evidence specific to collateral behavior changes and seems consistent with the intent of Reichow et al.'s recommendations.

Consistent with recommendations regarding effect size estimates for synthesis of SCD studies, three different nonparametric effect size estimates were calculated (Kratowill et al., 2013; Maggin & Odom, 2014).

Specifically, the percentage of nonoverlapping data (PND), improvement rate difference (IRD), and nonoverlap of all pairs (NAP) were calculated for all targeted and collateral behaviors (Parker & Vannest, 2009; Parker, Vannest, & Brown, 2009; Scruggs, Mastropieri, & Casto, 1987). SCD graphs were prepared for analysis by manually extracting the data from each study and saving the raw data into an Excel file. For multiple baseline, multiple probe, and reversal designs, all adjacent AB series (i.e. the intervention phase and the preceding baseline) were contrasted (Maggin, O'Keefe, & Johnson, 2011). For multi-element designs, data between the treatment and comparison condition were contrasted to determine an effect size estimate (Maggin et al., 2011). No alternating treatment designs in the included studies utilized more than two conditions.

PND was selected because it has been in use longer than other options which enables comparison to a larger portion of previous research (Campbell, 2013). Further, PND is the most commonly utilized effect size estimate in synthesis of SCDs (Maggin et al., 2011). To calculate PND, the number of data points in the intervention phase that exceed the highest baseline point is divided by the total number of data points in the intervention phase (Scruggs et al., 1987). PND effect size estimates range from 0% to 100% and were interpreted using the criteria outlined by Scruggs and Mastropieri (1998) wherein a PND value greater than 90% suggests a highly effective intervention, values between 70.1% and 90% a moderate effect, and values below 70% a low effect.

IRD represents the difference in improvement rate between baseline and intervention phases (Parker et al., 2009). It is highly correlated with Phi and mathematically equivalent to risk difference, a widely-used effect size estimate in medical research (Parker et al., 2009). Data points in intervention phases which exceed all baseline points are considered improved. IRD values range from 0 to 1 and those above .70 suggest a large effect, .50 to .70 a moderate effect, and values below .50 a small or questionable effect (Parker et al., 2009). NAP compares each data point from baseline and intervention in a pairwise fashion to determine a complete non-overlap index and is conceptualized as the percentage of data that improves across adjacent phases (Parker & Vannest, 2009). NAP values range from .50 to 1 and values of at least .93 suggest a large effect, .66 to .92 a moderate effect, and a small effect when at or below .65. Both IRD and NAP were calculated using the online calculator developed by Vannest, Parker, and Gonen (2011).

For studies utilizing group designs, Cohen's *d* was calculated for each reported variable using means and standard deviations (Cohen, 1988). Cohen's *d* is defined

as the standardized difference between group means and is common in synthesis of group design studies (Warner, 2013). Effect sizes of .20 and lower are considered small, values from .21 to .79 moderate, and values at or above .80 large (Warner, 2013).

Results

Table 1 summarizes the participant and intervention characteristics, effect size estimates (for SCDs) and Cohen's d (for group design studies), and methodological rigor of the 46 studies included in this systematic review. Table 2 provides average effect sizes for each collateral behavior and displays the variety of intervention and target behavior combinations that resulted in specific collateral outcomes. The narrative results that follow provide detail and summary to supplement Tables 1 and 2.

Participant characteristics

A total of 206 children (166 male) with ASD, ranging in age from 2;0 to 8;7 years ($M=4;2$), participated in the included studies. The majority of participants had characteristics consistent with Reichow and Volkmar's (2010) description of lower functioning ($n=94$ across 29 studies), followed by medium functioning ($n=40$ across 23 studies), and higher functioning ($n=13$ across seven studies). Seventeen studies included participants from different functioning levels (e.g. Charlop & Trasowech, 1991). These totals do not include participants from studies that did not provide enough detail to determine specific functioning level of included participants.

Intervention characteristics

A number of different intervention packages involving a variety of components were identified. Table 1 provides an exhaustive list but the most common examples include: naturalistic behavioral strategies (e.g. following the child's lead, use of natural reinforcers), prompting and reinforcement, script training, and motivating operation manipulation. Interventions were delivered in clinical settings ($n=18$), school settings ($n=10$), homes ($n=10$), and distraction-free locations in applied settings (e.g. empty rooms at a school; $n=10$). One study did not report the location of the intervention. Interventionists included researchers or graduate-level trained therapists ($n=32$), teachers ($n=6$), parents ($n=7$), and peers ($n=2$). Four studies utilized multiple intervention agents and five implemented the intervention across multiple settings. The total duration of intervention ranged from one to 400 hours ($M=21$ hrs.) with the majority involving fewer than ten hours

(median = 4 hrs.). Ten studies did not report intervention duration.

Target behaviors

Target behaviors in SCD studies included vocal utterances and requests ($n=8$), social language (e.g. initiations, social question-asking, bids for joint attention; $n=7$), stereotypy ($n=7$), joint attention ($n=4$), functional play ($n=3$), social play ($n=3$), expressive and receptive identification of stimuli ($n=3$), academic skills ($n=3$), daily-living tasks ($n=2$), challenging behavior, compliance, observation of conditioned reinforcers, imitation, and social interaction ($n=1$ each; see Table 1). Interventions investigated in group design studies targeted joint attention and symbolic play (Kasari et al., 2006) and prelinguistic joint attention acts and functional communication (Yoder & Stone, 2006b). Table 1 reports target behavior effect size estimates.

Collateral outcomes

Fourteen different collateral effects were identified across studies. Table 2 organizes studies in three groups according to how collateral effects align with the DSM-5's diagnostic criteria for ASD with social communication skills in group one and restricted/repetitive behaviors in group two. The third group included studies with collateral effects in domains not directly related to the DSM-5's diagnostic criteria (e.g. challenging behavior). For each specific collateral effect, Table 2 also provides the mean PND, IRD, and NAP scores across SCD studies or Cohen's d for group studies involving each collateral effect (column 1) and lists all the combinations of interventions and target behaviors associated with each collateral effect (column 2). For example, the first entry in Table 2 indicates that a total of four studies reported improved identification of stimuli (a skill related to receptive language) as a collateral effect and .79, .81, .88 are the mean PND, IRD, and NAP scores (respectively) for that collateral effect across the four studies. The intervention procedures and target behaviors involved in those four studies were Discrete Trial Teaching (DTT) targeting expressive and receptive language ($n=2$), matrix training targeting spelling words ($n=1$), and naturalistic behavioral strategies targeting question-asking ($n=1$).

Collateral effects related to communication and social interaction skills were reported in 34 cases. In terms of communication, average effect size estimates for collateral behavior changes in SCD studies ranged from low to moderate and included increased verbal utterances ($n=8$); language variability or vocabulary

Table 2. Summary of collateral effects by intervention and targeted skill.

Collateral effects related to communication and social skills	Intervention × Targeted Skill (n of Studies)
Identification of stimuli (.79, .81, .88) ^a	Discrete trial teaching × Expressive/ receptive language (2) Matrix training × Spelling (1) Naturalistic behavioral strategies × Social question-asking (1)
Utterances (.46, .46, .68) (Cohen's <i>d</i> range = .50–.71)	Abolishing operation component × Challenging behavior (1) Computer software × Language targets on computer (1) Naturalistic behavioral strategies × Imitation (1) Naturalistic behavioral strategies × Joint attention (2) Naturalistic behavioral strategies × Play (1) Picture Exchange Communication System × Requests (1) Reinforcement × Compliance (1) Response Interruption and Redirection/stimulus control × Stereotypy (3)
Varied or Novel Language (.44, .51, .76)	Naturalistic behavioral strategies × Social question-asking (1) Prompting and reinforcement × Social language (1) Script training × Social language (5)
Eye contact/ orienting (.87, .81, .91)	Computer software × Language targets on computer (1) Mand training (establishing operation contrived) × Requests (1) Naturalistic behavioral strategies with embedded social reinforcers × Utterances (2)
Joint attention (.60, .69, .83) (Cohen's <i>d</i> = .66)	Picture Exchange Communication System × Requests (1) Perseverative interests used in games × Social play (1) Perseverative interest in naturalistic behavioral strategies × Utterances (1) Naturalistic behavioral strategies × Imitation (1) Naturalistic behavioral strategies × Joint attention responses (1) Naturalistic behavioral strategies × Play (1)
Social Interaction (.76, .79, .94)	Naturalistic behavioral strategies with embedded social reinforcers × Utterances (1) Peer-mediated instruction × Reading comprehension (1)
Collateral effects related to restrictive repetitive behavior	Intervention × Targeted Skill (n of Studies)
Stereotypy (.67, .72, .87)	Peer mediated instruction × Social interaction (1) Perseverative interests used in games × Social play (1) Prompting and reinforcement × Play (2) Visual work system × On-task behavior in academics (1) Intervention × Targeted Skill (n of Studies) Reinforcement × Compliance (1)
Other collateral behaviors	
Attending (.84, .84, .93)	Mand training (establishing operation contrived) × Requests (1)
Compliance (.66, .48, .72)	
Challenging behavior (.71, .77, .87)	Abolishing operation component × Stereotypy (2) Computer software × Language targets on computer (1)

(continued)

Table 2. (continued)

Collateral effects related to communication and social skills	Intervention × Targeted Skill (n of Studies)
Functional or pretend play (.44, .49, .74)	Mand training (establishing operation contrived) × Requests (1) Naturalistic behavioral strategies × Utterances (2) Picture Exchange Communication System × Requests (1) Visual work system × Daily-living skills (2) Abolishing operation component × Stereotypy (2) Naturalistic behavioral strategies × Imitation (1) Punishment × Stereotypy (1) Script training × Social play (1) Video modeling × Social play (1) Naturalistic behavioral strategies × Joint attention (1)
Imitation (NR, NR, NR)	Conditioning new reinforcers × Observing of conditioned reinforcers (1)
Matching (1, 1, 1)	Non-contingent music × Stereotypy (1)
On-task behavior (.66, .33, .66)	

^aEffect Size Estimate Averages (PND, IRD, and NAP).

($n = 7$); and expressive and receptive identification of stimuli ($n = 4$). Cohen's d ranged from .50 to .71 for the three group design studies reporting a collateral increase in verbal utterances, indicating a moderate effect. A variety of intervention packages and components occasioned those collateral effects. Although not specifically listed in Table 2, all of the intervention packages involved some form of systematic prompting and reinforcement. The next most common intervention characteristic associated with collateral gains in communication involved naturalistic reinforcement contingencies delivered in developmentally appropriate natural contexts ($n = 6$); for example, naturalistic reinforcement contingencies embedded in play or daily routines (e.g. Koegel et al., 2014). Script training to target scripted spoken language resulted in collateral improvements in unscripted spoken language in five studies (e.g. Ledbetter-Cho et al., 2015). Two studies used Response Interruption and Redirection (RIRD) and one used a stimulus control procedure to reduce a targeted form of stereotypy. Collateral improvements in verbal utterances were reported in seven out of seven children in those studies (e.g. Ahearn et al., 2007).

In regards to social skills, average collateral effect size estimates for SCD studies ranged from low to moderate (see Table 2) and included improved joint attention ($n = 4$); eye contact or orientation toward a social partner ($n = 4$); and social interaction ($n = 2$). The two group design studies reported a moderate increase in joint attention, with Cohen's d equaling .66. In terms of commonalities across interventions that reported collateral effects in social skills, prompting and reinforcement were components of the intervention packages in all 12 cases. Seven targeted language skills (e.g. Vismara & Lyons, 2007) and six studies specifically described naturalistic behavioral strategies (e.g. Kasari et al., 2006; Koegel et al., 2009). Two studies incorporated participants' perseverative interests into intervention procedures targeting play or language skills and reported collateral improvement in joint attention (Baker, 2000; Vismara & Lyons, 2007). Finally, one study used peer-mediated instruction to target reading comprehension and reported an improvement in social interaction (Kamps et al., 1994).

Regarding restrictive and repetitive patterns of behavior and interests, a collateral decrease in some form of stereotypy (i.e. motor or vocal) was reported in five studies with a low mean effect size estimate. In all five studies, intervention involved some form of systematic prompting and reinforcement, one involved peer-mediated instruction, one incorporated participants' perseverative interests into games, two implemented prompting and reinforcement, and one utilized visual work systems. The most common target behavior associated with a collateral decrease in stereotypy was some

form of play ($n=3$; e.g. Lang et al., 2014). One study targeted social interaction (Lee & Odom, 1996) and one on-task behavior during academics (Bennett et al., 2011).

Collateral effects in domains and behaviors not explicitly required in the DSM-5's diagnostic criteria for ASD included challenging behavior ($n=9$), play ($n=6$), attending, compliance, imitation, matching, and on-task behavior ($n=1$ each). Average effect size estimates ranged from low to high, with one study not reporting enough information for calculation (Whalen et al., 2006). All but three studies included prompting and/or reinforcement; specifically: (a) Koegel et al. (1974) used punishment to decrease stereotypy and found an increase in appropriate play; (b) MacDonald et al. (2009) implemented video modeling to teach play and observed an increase in verbalizations that were not modeled in the video; and (c) Lanovaz et al. (2014) provided noncontingent access to music in an effort to decrease vocal stereotypy and reported improved on-task behavior. A decrease in stereotypy was the most common target behavior associated with collateral improvements in this group of studies ($n=6$) followed by requesting ($n=3$) and play ($n=2$). Naturalistic behavioral strategies were associated with four of these collateral effects (e.g. Gianoumis et al., 2012; Ingersoll & Schreibman, 2006).

Research designs and rigor

Interventions were evaluated in randomized controlled trials (RCT) in four studies (Kasari et al., 2006, 2008; Yoder & Stone, 2006a, 2006b) and the remainder were SCD. With regard to the target behaviors across studies, 18 (32%) were rated as having strong methodological rigor. Twenty-five target behaviors (44%) were rated as adequate. Of these variables, most received a rating of adequate due to lower scores on visual analysis (i.e. stability of the data, overlap between adjacent phases, and a lack of shift between conditions; $n=20$). Four of these variables received an adequate rating due to a lack of secondary quality indicators and one did not provide a sufficient description of participant characteristics. Fourteen target behaviors (24%) received ratings of weak methodological rigor due to an insufficient number of data points in baseline and/or intervention phases ($n=10$), inadequate stability in the data ($n=3$), or an absence of secondary quality indicators ($n=1$).

Each collateral effect across studies was also coded for rigor, with some studies receiving different ratings on different collateral behaviors (e.g. Baker, 2000). Ten collateral behaviors (14%) received strong ratings of research quality. Thirty-five collateral behaviors (48%) were rated as adequate due to overlap and

stability of data ($n=31$), insufficient number of baseline data points ($n=2$), or lack of secondary quality indicators ($n=2$). The twenty-eight remaining collateral effects (38%) received ratings of weak as a result of excessive overlap or variability in data ($n=15$), reporting averages across phases precluding visual analysis ($n=7$), insufficient number of baseline data points ($n=3$), or lack of secondary quality indicators ($n=3$).

Discussion

This systematic review of 46 intervention studies resulted in the identification of 14 general collateral effects (Table 2). The most common collateral effects involved behaviors directly related to ASD diagnostic criteria. Specifically, in terms of social communication skills, the following collateral increases were reported: (a) spoken utterances; (b) novel and more varied language; (c) joint attention, eye contact, and orienting toward a communication partner; (d) social interactions; and (e) receptive and expressive identification of stimuli. In regards to the amelioration of restrictive and repetitive behaviors, collateral decreases were found in motor and vocal stereotypic behavior. Improvements in skills not directly aligned with ASD diagnostic criteria were also reported (e.g. decreased challenging behavior). Overall, this systematic review identified a wide range of collateral effects across multiple domains of functioning and supports conclusions of previous reviews focused on specific intervention packages or target behaviors (e.g. Ganz et al., 2012; Lanovaz et al., 2013; Verschuur et al., 2014; White et al., 2011).

The finding that 206 children across 46 studies evidenced a collateral change in behavior suggests that these effects may not be uncommon. Further, given that only one participant's collateral behavior change was undesirable (i.e. Cook et al., 2014), collateral benefits appear to be more common than undesirable collateral side effects. However, the commonality of beneficial collateral effects should be considered cautiously because the potential for a collateral effect is not always considered when planning intervention research and may often go unmeasured. Similarly, researchers may be less likely to report efforts aimed at measuring potential collateral effects when no collateral changes are detected. Finally, the absence of consistent terminology to describe collateral effects complicates database searches: for example, we counted thirty-six different terms referring to collateral effects across included studies (list of terms available on request). Considered in tandem with differences in participant characteristics and intervention procedures across studies, these factors preclude calculating the probability of collateral effects for a given scenario

with sufficient certainty. Although the exact factors contributing to the probability of a collateral effect cannot be determined, notable trends across the included studies did emerge that suggest directions for future research and considerations for practitioners.

Within-study effect size estimates for target and collateral behaviors can be compared in 26 SCD studies that provided session-by-session data for all target and collateral behaviors (Table 1). In 20 of those studies (77%), every target behavior effect size estimate was larger than every collateral behavior effect size estimate in the same study and, in one of the remaining six studies, the effect size estimates for target and collateral behavior changes were equivalent (i.e. Krantz et al., 1993). The finding that target behavior effect size estimates tended to be larger suggests that interventions should include components that directly target the highest treatment priorities when possible. However, if intervention for a target behavior is unavailable, ineffective, or inefficient, it may be beneficial to initiate an intervention that targets a different behavior and/or include components that have been demonstrated to produce a collateral behavior change consistent with the goals of the initial focused intervention. For example, peer-mediated instruction targeting academics has produced collateral improvements in social skills and could be used to improve both academics and supplement a concurrent intervention targeting social skill deficits in cases where a child does not have access to a quality social skills intervention or the acquisition of targeted social skills has been slow (e.g. Kamps et al., 1994).

Many studies reporting collateral skill increases involved behaviors that were occasionally emitted prior to intervention. For example, children that produced at least a few vocal utterances prior to intervention appear to be more likely to experience collateral increases in utterances following intervention targeting joint attention than children who did not (e.g. Ingersoll & Schriebman, 2006). There were 129 demonstrations of collateral increases across studies that measured collateral behaviors in baseline sessions (e.g. a study with three participants that measured two potential collateral behaviors per participant could have up to six demonstrations of a collateral effect). Of the 129 collateral increases demonstrated across studies, 97 (75%) had two or more baseline sessions in which the collateral behavior occurred. This suggests that collateral increases may be more likely when there is a performance deficit as opposed to a skill deficit. Specifically, when a child has acquired a skill but does not demonstrate the skill at desired levels because stimulus control or motivation is insufficient (performance deficit), collateral increases in that skill may be more probable than in cases where the skill has not yet been acquired

and is therefore absent in baseline (skill deficit). It is important to note that the nonoccurrence of a skill across baseline sessions does not necessarily indicate that there is a skill deficit and not a performance deficit. However, the observation that the majority of collateral skills were demonstrated, at least to some degree, by participants prior to intervention suggests future research considering the potential influence of preexisting skill levels on collateral skill increases may be worthwhile.

In three studies, a collateral increase in a behavior was demonstrated despite an absence of evidence of the skill in baseline (i.e. potential skill deficit). The collateral behaviors in those studies were notably similar to the target behaviors. Specifically, Wichnick, Verner, Pyrttek, et al. (2010) targeted scripted responses to social initiations using script training and reported a collateral increase in novel (unscripted) social responses. Koegel et al. (2014) used a naturalistic behavioral intervention package to teach children to ask specific target questions and reported a collateral increase in question forms not targeted by intervention. Pollard et al. (2012) targeted scripted bids for joint attention using script training and reported a collateral increase in types of joint attention bids that were not directly scripted during intervention. In all three of these studies, the collateral behaviors and target behaviors likely shared the same operant function and involved similar discriminative stimuli (i.e. stimuli that precede the responses and signal potential reinforcement), suggesting that response generalization was likely the mechanism responsible for the collateral gains (Cooper et al., 2007; Kazdin, 1994). Response generalization may be facilitated by reinforcing variability in the topography of a target behavior or novel combinations of previously acquired behaviors (Kinney et al., 2003; Lee, Sturmey, & Fields, 2007; Pauwels et al., 2015). Studies utilizing research designs specifically arranged to test whether specific strategies (e.g. matrix training, multiple exemplars) are responsible for collateral effects may be especially useful (e.g. Kinney et al., 2003; Lang et al., 2014).

Collateral effects occurring in different domains and/or involving different operant functions or discriminative stimuli than the target behaviors are more consistent with the concepts of pivotal response and behavioral cusp than response generalization. Target behaviors involving play skills, communication/language, joint attention, and stereotypy were the most common among studies reporting collateral effects that do not meet definitions of response generalization (Stewart et al., 2013). In regards to joint attention, language, and play, this finding is consistent with a large corpus of previous research and highlights the potential bidirectional nature of interactions between these

variables. Specifically, interventions that target joint attention and/or play have reported collateral increases in language while interventions targeting language have occasioned collateral increases in play and joint attention (e.g. Baker, 2000; Kasari et al., 2006, 2008; Vismara & Lyons, 2007; Whalen et al., 2006; Yoder & Stone, 2006a). These collateral effects buttress conclusions of previous research linking joint attention and play to the emergence of language in children of typical development (e.g. Charman et al., 2000; Kuhn, Willoughby, Wilbourn, Vernon-Feagans, & Blair, 2014) and research suggesting that targeting developmentally appropriate behaviors (e.g. play in early childhood) may facilitate more efficient skill acquisition (e.g. Lifter et al., 1993, 2005). Further, because language ability predicts academic achievement, socialization, and executive functioning (e.g. Bono, Daley, & Sigman, 2004; Charman et al., 2003; Hart & Risely, 1995; Mundy, Sigman, & Kasari, 1990), it is not surprising that improved language can positively influence a wide range of additional variables including social interaction and challenging behavior (e.g. Charlop-Christy et al., 2002; Gianoumis et al., 2012; Koegel et al., 2009).

Targeted decreases in stereotypy were also associated with collateral improvements across a range of behaviors involving play, language, challenging behavior, and on-task behavior. Lanovaz et al.'s (2013) review noted that interventions aimed at reducing stereotypy should provide access to alternative activities or directly prompt appropriate replacement behaviors (e.g. play) to reduce the likelihood of undesirable collateral increases in another form of stereotypy or challenging behavior. Consistent with Lanovaz et al.'s recommendation, we found that a collateral increase in play was often reported following interventions targeting stereotypy and, conversely, a collateral decrease in stereotypy was often found following interventions targeting play (e.g. Baker, 2000; Koegel et al., 1974; Lang et al., 2009, 2010, 2014; Nuzzolo-Gomez et al., 2002). Those studies hypothesized that play and stereotypy may, in some cases, be maintained by similar operant functions (e.g. automatic reinforcement) which could facilitate collateral effects. However, the operant functions of the specific play and stereotypic behaviors were not directly assessed, and the extent to which a shared operant function between play and stereotypy contributes to the emergence of collateral effects warrants additional research.

Behavioral intervention components (e.g. prompting and reinforcement) embedded in naturalistic routines and activities constituted the most common intervention packages associated with collateral behavior change. Naturalistic behavioral intervention packages (e.g. PRT, Incidental Teaching) often involve parents or teachers as interventionists and are implemented in

applied settings, which may facilitate collateral effects by helping to ensure intervention is delivered for more hours per day and across multiple environments (e.g. Vernon et al., 2012). Alternatively, it is possible that measuring and reporting collateral effects is simply more common in studies using these procedures and that similar collateral effects arise from other intervention approaches but simply go unmeasured. Future research comparing collateral effects resulting from naturalistic behavioral intervention packages to behavioral interventions involving more contrived stimuli (e.g. DTT) could help elucidate intervention characteristics that contribute to collateral effects.

In terms of participant characteristics that corresponded with collateral effects, participant summaries provided in Table 1 reveal that the majority of children across studies (64%) had very limited vocal communication skills and/or an IQ below 55 (i.e. lower functioning per Reichow & Volkmer, 2010) and only 9% had age-appropriate vocal communication and an average or above-average IQ (higher functioning). Although it is possible that the potential for collateral effects decreases as participant functioning level increases, the studies included in the current review consisted of younger participants who may have been more likely to exhibit severe symptoms. Future research involving controls for level of functioning could identify interactions between participant characteristics, intervention procedures, and target skills that contribute to collateral effects. For example, it is possible that the intervention procedures used more often with participants that are lower functioning facilitate collateral gains (e.g. establishing a context for joint attention through naturalistic strategies while targeting play; Kasari et al., 2006) and/or that developmentally appropriate target skills are more likely to produce collateral effects than target skills that are not properly aligned with participant functioning level (e.g. Lifter et al., 2005).

Limitations and future research

In order to consider a larger sample of studies, we chose not to exclude studies based on number of participants, specific research designs, target behaviors, or intervention characteristics. Although this allowed for a broad-based consideration of the literature, it also precluded use of the fine-grained meta-analytic procedures potentially capable of determining the extent to which specific factors influenced collateral effects (e.g. Shadish, Hedges, & Pustejovsky, 2014). Additionally, the effect size estimates calculated from SCDs quantify the degree of overlap between measurements of dependent variables across baseline and intervention phases but do not necessarily reflect the magnitude of behavior change. Future research reviews should attempt to

identify potential moderators of collateral behavior change; however, that endeavor will require development of a novel or refined approach to calculating standardized effect sizes that can be utilized across a wider range of SCD variants (Pustejovsky & Ferron, 2017).

Regardless, in most cases, the included studies were rated as having strong or adequate methodological rigor, providing some certainty that effects reported in the included studies were not the result of maturation, concomitant intervention, measurement error, or other similar confounds. It is likely that additional focused intervention studies have produced collateral behavior changes that have not been measured or reported in the literature. Future research that measures collateral behavior changes throughout all phases of the study would provide additional insight into the nature of such behavior change. It is important to note that the majority of studies focused experimental controls on target behaviors and not collateral effects, resulting in higher ratings of research rigor for targeted behaviors (Reichow et al., 2008). Future research, could specifically tailor controls to ensure a higher degree of certainty regarding collateral effects and should consider addressing research questions regarding the mechanism of action for collateral effects directly. For example, the majority of reviewed studies did not involve design features or controls directed at testing hypothesized mechanisms of action for collateral behavior changes (e.g. recombinative generalization). Research illuminating the cause of specific untargeted behavior changes would better inform intervention creation and delivery.

Acknowledgements

The authors would like to thank Dr. James Pustejovsky for sharing his suggestions and expertise regarding effect sizes for single-case design studies.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The research described in this article was supported in part by Grant H325H140001 from the Office of Special Education Programs, U.S. Department of Education. Nothing in the article necessarily reflects the positions or policies of the federal government, and no official endorsement by it should be inferred.

References

- Ahearn, W. H., Clark, K. M., MacDonald, R. P. F., & Chung, B. I. (2007). Assessing and treating vocal

stereotypy in children with autism. *Journal of Applied Behavior Analysis, 40*, 263–275. doi: 10.1901/jaba.2007.30-06

Ahrens, E. N., Lerman, D. C., Kodak, T., Worsdell, A. S., & Keegan, C. (2011). Further evaluation of response interruption and redirection as treatment for stereotypy. *Journal of Applied Behavior Analysis, 44*, 95–108. doi: 10.1901/jaba.2011.44-95

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

Baer, D. M., Wolf, M. M., & Risley, T. R. (1987). Some still-current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis, 20*, 313–327. doi: 10.1901/jaba.1987.20-313

Baker, M. J. (2000). Incorporating the thematic ritualistic behaviors of children with autism into games: Increasing social play interactions with siblings. *Journal of Positive Behavior Interventions, 2*, 66–84. doi: 10.1177/109830070000200201

Bennett, K., Reichow, B., & Wolery, M. (2011). Effects of structured teaching on the behavior of young children with disabilities. *Focus on Autism and Other Developmental Disabilities, 26*, 143–152. doi: 10.1177/1088357611405040

Bono, M. A., Daley, T., & Sigman, M. (2004). Relations among joint attention, amount of intervention and language gain in autism. *Journal of Autism & Developmental Disorders, 34*, 495–505. doi: 10.1007/s10803-004-2545-x

Bosch, S., & Fuqua, R. W. (2001). Behavioral cusps: A model for selecting target behaviors. *Journal of Applied Behavior Analysis, 34*, 123–125. doi: 10.1901/jaba.2001.34-123

Cadogan, S., & McCrimmon, A. W. (2015). Pivotal response treatment for children with autism spectrum disorder: A systematic review of research quality. *Developmental Neurorehabilitation, 18*, 137–144. doi: 10.3109/17518423.2013.845615

Campbell, J. M. (2013). Commentary on PND at 25. *Remedial and Special Education, 34*, 20–25. doi: 10.1177/0741932512454725

Carr, E. G., & Durand, V. M. (1985). Reducing behavior problems through functional communication training. *Journal of Applied Behavior Analysis, 18*, 111–126. doi: 10.1901/jaba.1985.18-111

Celiberti, D. A., Bobo, H. E., Kelly, K. S., Harris, S. L., & Handleman, J. S. (1997). The differential and temporal effects of antecedent exercise on the self-stimulatory behavior of a child with autism. *Research in Developmental Disabilities, 18*, 139–150. doi: 10.1016/S0891-4222(96)00032-7

Chamak, B., & Bonniau, B. (2016). Trajectories, long-term outcomes and family experiences of 76 adults with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 46*, 1084–1095. doi: 10.1007/s10803-015-2656-6

Charlop, M. H., & Trasowech, J. E. (1991). Increasing autistic children's daily spontaneous speech. *Journal of Applied Behavior Analysis, 24*, 747–761. doi: 10.1901/jaba.1991.24-747

Charlop-Christy, M. H., Carpenter, M., Le, L., LeBlanc, L. A., & Kellet, K. (2002). Using the picture exchange

- communication system (PECS) with children with autism: Assessment of PECS acquisition, speech, social-communicative behavior, and problem behavior. *Journal of Applied Behavior Analysis*, *35*, 213–231. doi: 10.1901/jaba.2002.35-231
- Charman, T., Baron-Cohen, S., Swettenham, J., Baird, G., Cox, A., & Drew, A. (2000). Testing joint attention, imitation, and play as infancy precursors to language and theory of mind. *Cognitive Development*, *15*, 481–498. doi: 10.1016/S0885-2014(01)00037-5
- Charman, T., Baron-Cohen, S., Swettenham, J., Baird, G., Drew, A., & Cox, A. (2003). Predicting language outcome in infants with autism and pervasive developmental disorder. *International Journal of Language & Communication Disorders*, *38*, 265–285. doi: 10.1080/13682031000104830
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.
- Cook, J. L., Rapp, J. T., Gomes, L. A., Frazer, T. J., & Lindblad, T. L. (2014). Effects of verbal reprimands on targeted and untargeted stereotypy. *Behavioral Interventions*, *29*, 106–124. doi: 10.1002/bin.1378
- Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). *Applied behavior analysis* (2nd ed.). Upper Saddle River, NJ: Pearson Prentice Hall.
- Dawson, G., Jones, E. J. H., Merkle, K., Venema, K., Lowy, R., Faja, S., . . . Webb, S. J. (2012). Early behavioral intervention is associated with normalized brain activity in young children with autism. *Journal of the American Academy of Child and Adolescent Psychiatry*, *51*, 1150–1159. doi: 10.1016/j.jaac.2012.08.018
- Davis, T. N., O'Reilly, M. F., Kang, S., Rispoli, M., Lang, R. B., Machalicek, W., . . . Sigafos, J. (2009). Impact of pre-session access to toys maintaining challenging behavior on functional communication training: A single case study. *Journal of Developmental & Physical Disabilities*, *21*, 515–521. doi: 10.1007/s10882-009-9158-4
- Delfs, C. H., Conine, D. E., Frampton, S. E., Shillingsburg, M. A., & Robinson, H. C. (2014). Evaluation of the efficiency of listener and tact instruction for children with autism. *Journal of Applied Behavior Analysis*, *47*, 793–809. doi: 10.1002/jaba.166
- Ganz, J. B., Davis, J. L., Lund, E. M., Goodwyn, F. D., & Simpson, R. L. (2012). Meta-analysis of PECS with individuals with ASD: Investigation of targeted versus non-targeted outcomes, participant characteristics, and implementation phase. *Research in Developmental Disabilities: A Multidisciplinary Journal*, *33*, 406–418.
- Gianoumis, S., Seiverling, L., & Sturmey, P. (2012). The effects of behavioral skills training on correct teacher implementation of natural language paradigm teaching skills and child behavior. *Behavioral Interventions*, *27*, 57–74. doi: 10.1002/bin.1334
- Goldstein, H., & Cisar, C. L. (1992). Promoting interaction during sociodramatic play: Teaching scripts to typical preschoolers and classmates with disabilities. *Journal of Applied Behavior Analysis*, *25*, 265–280. doi: 10.1901/jaba.1992.25-265
- Greer, R. D., & Han, H. S.-A. (2015). Establishment of conditioned reinforcement for visual observing and the emergence of generalized visual-identity matching. *Behavioral Development Bulletin*, *20*, 227–252. doi: 10.1037/h0101316 227-252
- Hart, B., & Risely, T. R. (1995). *Meaningful difference in the everyday experiences of young American children*. Baltimore, MA: Paul Brookes.
- Henninger, N. A., & Taylor, J. L. (2012). Outcomes in adults with autism spectrum disorders: A historical perspective. *Autism*, *17*, 103–116. doi: 10.1177/1362361312441266
- Higgins, J. P. T., & Green, S. (2011). *Cochrane handbook for systematic reviews of interventions (Version 5.1.0)*. The Cochrane Collaboration, 2011. Retrieved from www.cochranehandbook.org
- Ingersoll, B., & Schreibman, L. (2006). Teaching reciprocal imitation skills to young children with autism using a naturalistic behavioral approach: Effects on language, pretend play, and joint attention. *Journal of Autism & Developmental Disorders*, *36*, 487–505. doi: 10.1007/s10803-006-0089-y
- Jacobson, J. W., & Mulick, J. A. (2000). System and cost research issues in treatments for people with autistic disorders. *Journal of Autism & Developmental Disorders*, *30*, 585–593. doi: 10.1023/A:1005691411255
- Kamps, D. M., Barbetta, P. M., Leonard, B. R., & Delquadri, J. (1994). Classwide peer tutoring: An integration strategy to improve reading skills and promote peer interactions among students with autism and general education peers. *Journal of Applied Behavior Analysis*, *27*, 49–61. doi: 10.1901/jaba.1994.27-49
- Karaaslan, O., Diken, I. H., & Mahoney, G. (2013). A randomized control study of responsive teaching with young Turkish children and their mothers. *Topics in Early Childhood Special Education*, *33*, 18–27. doi: 10.1177/0271121411429749
- Kasari, C., Freeman, S., & Paparella, T. (2006). Joint attention and symbolic play in young children with autism: A randomized controlled intervention study. *Journal of Child Psychology & Psychiatry*, *47*, 611–620. doi: 10.1111/j.1469-7610.2005.01567.x
- Kasari, C., Paparella, T., Freeman, S., & Jahromi, L. B. (2008). Language outcome in autism: Randomized comparison of joint attention and play interventions. *Journal of Consulting and Clinical Psychology*, *76*, 125–137. doi: 10.1037/0022-006X.76.1.125
- Kazdin, A. E. (1994). *Behavior modification in applied settings* (5th ed.). Pacific Grove, CA: Brooks-Cole.
- Kinney, E. M., Vedora, J., & Stromer, R. (2003). Computer-presented video models to teach generative spelling to a child with an autism spectrum disorder. *Journal of Positive Behavior Interventions*, *5*, 22–29. doi: 10.1177/10983007030050010301
- Koegel, L. K., Camarata, S. M., Valdez-Menchaca, M., & Koegel, R. L. (1998). Setting generalization of question-asking by children with autism. *American Journal on Mental Retardation*, *102*, 346–357. doi: 10.1352/0895-8017
- Koegel, R. L., Koegel, L. K., & Surratt, A. (1992). Language intervention and disruptive behavior in preschool children with Autism. *Journal of Autism and Developmental Disorders*, *22*, 141–153. doi: 10.1007/BF01058147

- Koegel, L. K., Ashbaugh, K., & Koegel, R. L. (2016). Pivotal response treatment. In R. Lang, T. Hancock & N. N. Singh (Eds.), *Early intervention for young children with autism*. New York, NY: Springer.
- Koegel, R. L., Bradshaw, J. L., Ashbaugh, K., & Koegel, L. K. (2014). Improving question-asking initiations in young children with autism using pivotal response treatment. *Journal of Autism & Developmental Disorders, 44*, 816–827. doi: 10.1007/s10803-013-1932-6
- Koegel, R. L., Firestone, P. B., Kramme, K. W., & Dunlap, G. (1974). Increasing spontaneous play by suppressing self-stimulation in autistic children. *Journal of Applied Behavior Analysis, 7*, 521–528. doi: 10.1901/jaba.1974.7-521
- Koegel, R. L., & Koegel, L. K. (2006). *Pivotal response treatments for autism: Communication, social, & academic development*. Baltimore, MD: Paul H. Brookes.
- Koegel, R. L., Koegel, L. K., & McNeerney, E. K. (2001). Pivotal areas in intervention for autism. *Journal of Clinical Child Psychology, 30*, 19–32. doi: 10.1207/S15374424JCCP3001_4
- Koegel, R. L., & Mentis, M. (1985). Motivation in childhood autism: Can they or won't they? *Journal of Child Psychology and Psychiatry, 26*, 185–191. doi: 10.1111/j.1469-7610.1985.tb02259.x
- Koegel, R. L., Vernon, T. Y., & Koegel, L. K. (2009). Improving social initiations in young children with autism using reinforcers with embedded social interactions. *Journal of Autism & Developmental Disorders, 39*, 1240–1251. doi: 10.1007/s10803-009-0732-5
- Koegel, R. L., & Wilhelm, H. (1973). Selective responding to the components of multiple visual cues by autistic children. *Journal of Experimental Child Psychology, 15*, 442–453. doi: 10.1016/0022-0965(73)90094-5
- Krantz, P. J., MacDuff, M. T., & McClannahan, L. E. (1993). Programming participation in family activities for children with autism: Parents' use of photographic activity schedules. *Journal of Applied Behavior Analysis, 26*, 137–138. doi: 10.1901/jaba.1993.26-137
- Kratochwill, T. R., Hitchcock, J. H., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., . . . Shadish, W. R. (2013). Single-case intervention research design standards. *Remedial and Special Education, 34*, 26–38. doi: 10.1177/0741932512452794
- Kuhn, L. J., Willoughby, M. T., Wilbourn, M. P., Vernon-Feagans, L., & Blair, C. B. (2014). Early communicative gestures prospectively predict language development and executive function in early childhood. *Child Development, 85*, 1898–1914. doi: 10.1111/cdev.12249
- Lang, R., Hancock, T. & Singh, N. N. (Eds.). (2016). *Early intervention for young children with autism*. New York, NY: Springer.
- Lang, R., Machalicek, W., Rispoli, M., O'Reilly, M., Sigafoos, J., Lancioni, G., . . . Didden, R. (2014). Play skills taught via behavioral intervention generalize, maintain, and persist in the absence of socially mediated reinforcement in children with autism. *Research in Autism Spectrum Disorders, 8*, 860–872. doi: 10.1016/j.rasd.2014.04.007
- Lang, R., O'Reilly, M., Sigafoos, J., Lancioni, G. E., Machalicek, W., Rispoli, M., & White, P. (2009). Enhancing the effectiveness of a play intervention by abolishing the reinforcing value of stereotypy: A pilot study. *Journal of Applied Behavior Analysis, 42*, 889–894. doi: 10.1901/jaba.2009.42-889
- Lang, R., O'Reilly, M., Sigafoos, J., Machalicek, W., Rispoli, M., Lancioni, G. E., . . . Fragale, C. (2010). The effects of an abolishing operation intervention component on play skills, challenging behavior, and stereotypy. *Behavior Modification, 34*, 267–289. doi: 10.1177/0145445510370713
- Lanovaz, M. J., Rapp, J. T., Maciw, I., Prigent-Pelletier, É., Dorion, C., Ferguson, S., & Saade, S. (2014). Effects of multiple interventions for reducing vocal stereotypy: Developing a sequential intervention model. *Research in Autism Spectrum Disorders, 8*, 529–545. doi: 10.1016/j.rasd.2014.01.009
- Lanovaz, M. J., Robertson, K. M., Soerono, K., & Watkins, N. (2013). Effects of reducing stereotypy on other behaviors: A systematic review. *Research in Autism Spectrum Disorders, 7*, 1234–1243. doi: 10.1016/j.rasd.2013.07.009
- Ledbetter-Cho, K., Lang, R., Watkins, L., & O'Reilly, M. (2015). Collateral Effects of Interventions Targeting Core Skill Deficits of Young Children with Autism: A Systematic Review. *PROSPERO International prospective register of systematic reviews*. CRD42015032301.
- Ledbetter-Cho, K., Lang, R., Davenport, K., Moore, M., Lee, A., Howell, A., . . . O'Reilly, M. (2015). Effects of script training on the peer-to-peer communication of children with autism spectrum disorder. *Journal of Applied Behavior Analysis, 48*, 785–799. doi: 10.1002/jaba.240
- Lee, R., Sturmey, P., & Fields, L. (2007). Schedule-induced and operant mechanisms that influence response variability: A review and implications for future investigations. *The Psychological Record, 57*, 429–455.
- Lee, S., & Odom, S. L. (1996). The relationship between stereotypic behavior and peer social interaction for children with severe disabilities. *Journal of the Association for Persons with Severe Handicaps, 21*, 88–95. doi: 10.1177/154079699602100204
- Lifter, K., Ellis, J., Cannon, B., & Anderson, S. R. (2005). Developmental specificity in targeting and teaching play activities to children with pervasive developmental disorders. *Journal of Early Intervention, 27*, 247–267.
- Lifter, K., Sulzer-Azaroff, B., Anderson, S. R., & Cowdery, G. E. (1993). Teaching play activities to preschool children with disabilities: The importance of developmental considerations. *Journal of Early Intervention, 17*, 139–159. doi: 10.1177/105381519301700206
- Lovaas, O. I., Koegel, R., Simmons, J. Q., & Long, J. S. (1973). Some generalization and follow-up measures on autistic children in behavior therapy. *Journal of Applied Behavior Analysis, 6*, 131–166. doi: 10.1901/jaba.1973.6-131
- McConnell, S. R. (2002). Interventions to facilitate social interaction for young children with autism: Review of available research and recommendations for educational intervention and future research. *Journal of Autism & Developmental Disorders, 32*, 351–372. doi: 10.1023/A:1020537805154
- MacDonald, R., Sacramone, S., Mansfield, R., Wiltz, K., & Ahearn, W. H. (2009). Using video modeling to teach reciprocal pretend play to children with autism. *Journal*

- of *Applied Behavior Analysis*, 42, 43–55. doi: 10.1901/jaba.2009.42-43
- McEvoy, M. A., Nordquist, V. M., Twardosz, S., Heckaman, K. A., Wehby, J. H., & Denny, R. K. (1988). Promoting autistic children's peer interaction in an integrated early childhood setting using affection activities. *Journal of Applied Behavior Analysis*, 21, 193–200. doi: 10.1901/jaba.1988.21-193
- Maggin, D. M., & Odom, S. L. (2014). Evaluating single-case research data for systematic review: A commentary for the special issue. *Journal of School Psychology*, 52, 237–241. doi: 10.1016/j.jsp.2014.01.002
- Maggin, D. M., O'Keefe, B. V., & Johnson, A. H. (2011). A quantitative synthesis of methodology in the meta-analysis of single-subject research for students with disabilities: 1985–2009. *Exceptionality*, 19, 109–135. doi: 10.1080/09362835.2011.565725
- Mannion, A., & Leader, G. (2013). Comorbidity in autism spectrum disorder: A literature review. *Research in Autism Spectrum Disorder*, 7, 1595–1616. doi: 10.1016/j.rasd.2013.09.006
- Matson, J. L., & Nebel-Schwalm, M. (2007). Assessing challenging behaviors in children with autism spectrum disorders: A review. *Research in Developmental Disabilities*, 28, 567–579. doi: 10.1016/j.ridd.2006.08.001
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine*, 151, 264–269. doi: 10.1371/journal.pmed.1000097
- Mundy, P., Sigman, M., & Kasari, C. (1990). A longitudinal study of joint attention and language development in autistic children. *Journal of Autism and Developmental Disorders*, 20, 115–128. doi: 10.1007/BF02206861
- Nuzzolo-Gomez, R., Leonard, M. A., Ortiz, E., Rivera, C. M., & Greer, R. D. (2002). Teaching children with autism to prefer books or toys over stereotypy or passivity. *Journal of Positive Behavior Interventions*, 4, 80–87. doi: 10.1177/109830070200400203
- O'Reilly, M., Falcomata, T. S., Kang, S., & Fragale, C. (2014). *Evidence-based treatment approaches for autism spectrum disorders: A review of the literature and recommendations* (DARS Autism Program Report). Retrieved from <http://www.dars.state.tx.us/autism/AutismReport.pdf>
- Parker, R. I., & Vannest, K. (2009). An improved effect size for single-case research: Nonoverlap of all pairs. *Behavior Therapy*, 40, 357–367. doi: 10.1016/j.beth.2008.10.006
- Parker, R. I., Vannest, K., & Brown, L. (2009). The improvement rate difference for single-case research. *Exceptional Children*, 75, 135–150. doi: 10.1177/001440290907500201
- Pauwels, A. A., Ahearn, W. H., & Cohen, S. J. (2015). Recombinative generalization of tacts through matrix training with individuals with autism spectrum disorder. *Analysis of Verbal Behavior*, 31, 200–214. doi: 10.1007/s40616-015-0038-y
- Pierce, K. L., & Schreibman, L. (1994). Teaching daily living skills to children with autism in unsupervised settings through pictorial self-management. *Journal of Applied Behavior Analysis*, 27, 471–481. doi: 10.1901/jaba.1994.27-471
- Pickard, K. E., & Ingersoll, B. R. (2016). Quality versus quantity: The role of socioeconomic status on parent-reported service knowledge, service use, unmet service needs, and barriers to service use. *Autism*, 20, 106–115. doi: 10.1177/1362361315569745
- Pickles, A., Couteur, A. L., Leadbitter, K., Salmone, E., Cole-Fletcher, R., Tobin, H., ... Green, J. (2016). Parent-mediated social communication therapy for young children with autism (PACT): Long-term follow-up of a randomised controlled trial. *Lancet*, 388, 2501–2509. doi: 10.1016/S0140-6736(16)31229-6
- Plavnick, J. B., & Ferreri, S. J. (2012). Collateral effects of mand training for children with autism. *Research in Autism Spectrum Disorders*, 6, 1366–1376. doi: 10.1016/j.rasd.2012.05.008
- Pollard, J. S., Betz, A. M., & Higbee, T. S. (2012). Script fading to promote unscripted bids for joint attention in children with autism. *Journal of Applied Behavior Analysis*, 45, 387–393. doi: 10.1901/jaba.2012.45-387
- Pustejovsky, J. E., & Ferron, J. M. (2017). Research synthesis and meta-analysis of single-case designs. In J. M. Kaufmann, D. P. Hallahan & P. C. Pullen (Eds.), *Handbook of special education* (2nd ed.). New York, NY: Routledge.
- Reichow, B., Barton, E. E., Boyd, B. A., & Hume, K. (2014). *Early intensive behavioral intervention (EIBI) for young children with autism spectrum disorders (ASD): A systematic review*. (Campbell Systematic Reviews 2014:9). Retrieved from <http://files.eric.ed.gov/libproxy.txstate.edu/fulltext/ED557969.pdf>
- Reichow, B., Barton, E. E., Sewell, J. N., Good, L., & Wolery, M. (2010). Effects of weighted vests on the engagement of children with developmental delays and autism. *Focus on Autism and Other Developmental Disabilities*, 25(1), 3–11. doi: 10.1177/1088357609353751
- Reichow, B., & Volkmar, F. (2010). Social skills interventions for individuals with autism: Evaluation for evidence-based practices within a best evidence synthesis framework. *Journal of Autism and Developmental Disorders*, 40, 149–166. doi: 10.1007/s10803-009-0842-0
- Reichow, B., Volkmar, F. R., & Cicchetti, D. V. (2008). Development of the evaluative method for evaluating and determining evidence-based practices in autism. *Journal of Autism and Developmental Disorders*, 38, 1311–1319. doi: 10.1007/s10803-007-0517-74
- Rocha, M. L., Schreibman, L., & Stahmer, A. C. (2007). Effectiveness of training parents to teach joint attention in children with autism. *Journal of Early Intervention*, 29, 154–173. doi: 10.1177/105381510702900207
- Rogers, S. J., Estes, A., Lord, C., Vismara, L., Winter, J., Fitzpatrick, A., ... Dawson, G. (2012). Effects of a brief Early Start Denver Model (ESDM)-based parent intervention on toddlers at risk for autism spectrum disorders: A randomized controlled trial. *Journal of the American Academy of Child & Adolescent Psychiatry*, 51, 1052–1065. doi: 10.1016/j.jaac.2012.08.003
- Ryberg, K. H. (2015). Evidence for the implementation of the early start Denver model for young children with autism spectrum disorder. *Journal of the American Psychiatric*

- Nurses Association*, 21, 327–337. doi: 10.1177/1078390315608165
- Sarokoff, R., Taylor, B., & Poulson, C. I. (2001). Teaching children with autism to engage in conversational exchanges: Script fading with embedded textual stimuli. *Journal of Applied Behavior Analysis*, 34, 81–84. doi: 10.1901/jaba.2001.34-81
- Schreibman, L., & Stahmer, A. C. (2014). A randomized trial comparison of the effects of verbal and pictorial naturalistic communication strategies on spoken language for young children with autism. *Journal of Autism and Developmental Disorders*, 44, 1244–1251. doi: 10.1007/s10803-013-1972-y
- Scruggs, T. E., & Mastropieri, M. A. (1998). Summarizing single-subject research: Issues and applications. *Behavior Modification*, 22, 221–242. doi: 10.1177/01454455980223001
- Scruggs, T. E., Mastropieri, M. A., & Casto, G. (1987). The quantitative synthesis of single-subject research methodology and validation. *Remedial and Special Education*, 8, 24–33. doi: 10.1177/074193258700800206
- Shadish, W. R., Hedges, L. V., & Pustejovsky, J. E. (2014). Analysis and meta-analysis of single-case designs with a standardized mean difference statistic: A primer and applications. *Journal of School Psychology*, 52, 123–147. doi: 10.1016/j.jsp.2013.11.005
- Siegel, M., & Beaulieu, A. (2012). Psychotropic medications in children with autism spectrum disorders: A systematic review and synthesis for evidence-based practice. *Journal of Autism & Developmental Disorders*, 42, 1592–1605. doi: 10.1007/s10803-011-1399-2
- Smith, G. J., McDougall, D., & Edelen-Smith, P. (2006). Behavioral cusps: A person-centered concept for establishing pivotal individual, family, and community behaviors and repertoires. *Focus on Autism & Other Developmental Disabilities*, 21, 223–229. doi: 10.1177/10883576060210040301
- Soutor, T. A., Houlihan, D., & Young, A. (1994). An examination of response covariation on the behavioral treatment of identical twin boys with multiple behavioral disorders. *Behavioral Interventions*, 9, 141–155. doi: 10.1002/bin.2360090302
- Stewart, I., McElwee, J., & Ming, S. (2013). Language generativity, response generalization, and derived relational responding. *The Analysis of Verbal Behavior*, 29, 137–155.
- Thomas, K. C., Ellis, A. R., McLaurin, C., Daniels, J., & Morrissey, J. P. (2007). Access to care for autism-related services. *Journal of Autism & Developmental Disorders*, 37, 1902–1912. doi: 10.1007/s10803-006-0323-7
- Thorp, D. M., Stahmer, A. C., & Schreibman, L. (1995). Effects of sociodramatic play training on children with autism. *Journal of Autism and Developmental Disorders*, 25, 265–282. doi: 10.1007/BF02179288
- Vannest, K., Parker, R. I., & Gonen, O. (2011). *Single case research: Web based calculators for SCR analysis (Version 1.0)*. College Station, TX: Texas A & M University [Webbased application]. Retrieved from singlecasere-search.org
- Vernon, T. Y., Koegel, R. L., Dauterman, H., & Stolen, K. (2012). An early social engagement intervention for young children with autism and their parents. *Journal of Autism & Developmental Disorders*, 42, 2702–2717. doi: 10.1007/s10803-012-1535-7
- Verschuur, R., Didden, R., Lang, R., Sigafoos, J., & Huskens, B. (2014). Pivotal response treatment for children with autism spectrum disorders: A systematic review. *Journal of Autism and Developmental Disorders*, 44, 34–61. doi: 10.1007/s40489-013-0008-z
- Virués-Ortega, J. (2010). Applied behavior analytic intervention for autism in early childhood: Meta-analysis, meta-regression and dose–response meta-analysis of multiple outcomes. *Clinical Psychology Review*, 30, 387–399. doi: 10.1016/j.cpr.2010.01.008
- Vismara, L. A., & Lyons, G. L. (2007). Using perseverative interests to elicit joint attention behaviors in young children with autism: Theoretical and clinical implications for understanding motivation. *Journal of Positive Behavior Interventions*, 9, 214–228. doi: 10.1177/10983007070090040401
- Vismara, L. A., & Rogers, S. J. (2010). Behavioral treatments in autism spectrum disorder: What do we know? *Annual Review of Clinical Psychology*, 6, 447–468. doi: 10.1146/annurev.clinpsy.121208.131151
- Vohra, R., Madhavan, S., Sambamoorthi, U., & Peter, C. S. (2014). Access to services, quality of care, family impact for children with autism, other developmental disabilities, and other mental health conditions. *Autism*, 18, 815–826. doi: 10.1177/1362361313512902
- Warner, R. M. (2013). *Applied statistics: From bivariate through multivariate techniques* (2nd.). Thousand Oaks, CA: SAGE Publications.
- Whalen, C., Schreibman, L., & Ingersoll, B. (2006). The collateral effects of joint attention training on social initiations, positive affect, imitation, and spontaneous speech for young children with autism. *Journal of Autism and Developmental Disorders*, 36, 655–664. doi: 10.1007/s10803-006-0108-z
- Whalen, C., Liden, L., Ingersoll, B., Dallaire, E., & Liden, S. (2006). Behavioral improvements associated with computer-assisted instruction for children with developmental disabilities. *The Journal of Speech and Language Pathology – Applied Behavior Analysis*, 1, 11–26. doi: 10.1037/h0100182
- Whalon, K. J., Conroy, M. A., Martinez, J. R., & Werch, B. L. (2015). School-based peer-related social competence interventions for children with autism spectrum disorder: A meta-analysis and descriptive review of single case research design studies. *Journal of Autism and Developmental Disorders*, 45, 1513–1531. doi: 10.1007/s10803-015-2373-1
- White, P. J., O'Reilly, M., Streusand, W., Levine, A., Sigafoos, J., Lancioni, G., . . . Aguilar, J. (2011). Best practices for teaching joint attention: A systematic review of the intervention literature. *Research in Autism Spectrum Disorders*, 5, 1283–1295.
- Wichnick, A. M., Vener, S. M., Keating, C., & Poulson, C. L. (2010). The effect of a script-fading procedure on unscripted social initiations and novel utterances among young children with autism. *Research in Autism Spectrum Disorders*, 4, 51–64. doi: 10.1016/j.rasd.2009.07.006

- Wichnick, A. M., Vener, S. M., Pyrtek, M., & Poulson, C. L. (2010). The effect of a script-fading procedure on responses to peer initiations among young children with autism. *Research in Autism Spectrum Disorders, 4*, 290–299. doi: 10.1016/j.rasd.2009.09.016
- Wynn, J. W., & Smith, T. (2003). Generalization between receptive and expressive language in young children with autism. *Behavioral Interventions, 18*, 245–266. doi: 10.1002/bin.142
- Yoder, P., & Stone, W. L. (2006a). A randomized comparison of the effect of two prelinguistic communication interventions on the acquisition of spoken communication in preschoolers with ASD. *Journal of Speech, Language, and Hearing Research, 49*, 698–711. doi: 1092-4388/06/4904-0698
- Yoder, P., & Stone, W. L. (2006b). Randomized comparison of two communication interventions for preschoolers with autism spectrum disorders. *Journal of Consulting and Clinical Psychology, 74*, 426–435. doi: 10.1037/0022-006X.74.3.426.