



Thinning Schedules of Reinforcement Following Functional Communication Training for Children with Intellectual and Developmental Disabilities: A Meta-analytic Review

Reem Muharib¹ · Fahad Alrasheed² · Jennifer Ninci³ · Virginia L. Walker⁴ · Ashley P. Voggt⁴

Published online: 27 August 2019

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

Abstract

Functional communication training (FCT) is an evidence-based practice used to mitigate challenging behavior by increasing functional communication skills. To increase the practicality and feasibility of FCT in natural settings, thinning schedules of reinforcement are typically programmed following FCT. In this review, we meta-analyzed 28 studies that incorporated a thinning schedule procedure following FCT for 51 children with intellectual and developmental disabilities ages 8 and younger. Using Tau-U, the results demonstrated overall moderate effect sizes for both challenging behavior and functional communication responses. Additionally, moderator analyses pertaining to participant characteristics, interventions, and study quality were conducted. Thinning procedures were most effective for children who had stronger communication repertoire. Implications for future research and practice are discussed.

Keywords Functional communication training · Reinforcement schedule thinning · Children with intellectual and developmental disabilities · Meta-analysis

Researchers have estimated 10–15% of children with intellectual and developmental disabilities (IDD) engage in challenging behavior such as aggression, self-injury, and property destruction (Emerson et al. 2001; Lowe et al. 2007). Challenging behavior can sometimes be attributed to limited functional communication skills (Park et al. 2012). One intervention used to mitigate challenging behavior by increasing appropriate communication skills is functional communication training (FCT; Carr and Durand 1985). In FCT, an individual is taught to use an appropriate functional communicative response (FCR) to access their wants and needs instead of engaging in challenging behavior (Carr and Durand 1985). Challenging behavior is typically placed on extinction, whereby the individual no longer receives

consequences that previously reinforced the challenging behavior (Hagopian et al. 2011). For example, a child may be taught to say, “I need to talk to you” as an alternative way to receive her mother’s attention rather than engaging in aggression.

Although FCT is an evidence-based practice (Muharib and Wood 2018; Wong et al. 2013) that has been effective in decreasing challenging behavior and increasing appropriate FCRs among children with IDD (e.g., Muharib et al. 2019), FCT may be unfeasible and impractical in natural settings because FCT requires caregivers/teachers to reinforce every FCR emitted by the child. In other words, the caregivers/teachers must give the child what they are asking for every time the child uses the newly learned communication response. This may raise an issue when the child uses the FCR too often. In this case, the caregiver/teacher may be unable to deliver what the child is requesting (e.g., due to being in a public place or requesting during an instructional period) which can cause the FCR to undergo extinction. When the child does not receive the reinforcer upon using the appropriate FCR, the child may return to engage in challenging behavior (Fisher et al. 2000; Hagopian et al. 2011).

✉ Reem Muharib
r_m806@txstate.edu

¹ Department of Curriculum and Instruction, Texas State University, 601 University Dr., San Marcos, TX 78666, USA

² Center for Behavioral Sciences, Irvine, CA, USA

³ University of Hawaii at Manoa, Honolulu, HI, USA

⁴ University of North Carolina at Charlotte, Charlotte, NC, USA

Schedule Thinning

To facilitate maintenance effects in natural settings, researchers have followed FCT with schedule thinning procedures to bring the FCR to a practical level while maintaining low levels of challenging behavior. Schedule thinning involves decreasing the rate or density of reinforcement until it meets the levels of reinforcement appropriate in the child's natural environment (Hagopian et al. 2011). Examples of thinning schedules of reinforcement following FCT include delay-to-reinforcement (e.g., Hanley et al. 2014), chained schedules of reinforcement (e.g., Falcomata et al. 2012a), multiple schedules of reinforcement (e.g., Greer et al. 2016), response restriction (e.g., Roane et al. 2004), and alternative activities which are often used to supplement multiple schedules of reinforcement (e.g., Hagopian et al. 2005). See Table 1 for definitions of these strategies.

There have been several reviews on the topic of FCT including those on the quality of the literature-base (e.g., Andzik et al. 2016; Durand and Moskowitz 2015), the quantified outcomes resulting from the intervention (e.g., Heath et al. 2015), and both quality and quantified outcomes (e.g., Chezan et al. 2018). However, few reviews have focused on FCT outcomes related to reinforcement schedule thinning to prevent reemergence of challenging behavior and promote natural schedules of reinforcement post-intervention. In their descriptive review of FCT research, Tiger et al. (2008) summarized strategies in reinforcement thinning, stressing the importance of this component for a socially valid intervention approach.

Hagopian et al. (2011) conducted a comprehensive review of FCT studies with reinforcement thinning strategies. These researchers described four methods of thinning that can be used with FCT. This study included descriptions and examples of each approach, analyses of their strengths and limitations, and brief literature summaries. In the literature summaries, the number of studies and applications conducted per approach was reported with the percentage of studies that used supplemental treatment components in the schedule thinning phase such as punishment or noncontingent reinforcement. In addition, participant diagnoses, target behavior typographies, and functions of challenging behavior among participants were reported. This review did not include effect sizes, standard analyses of study quality, or meta-analytic review components.

Quality of Literature Within FCT

Analyzing the quality of literature bases is an increasingly common practice in special education research because it allows for evaluations of internal validity within and across research studies. The use of standards to analyze quality also allows practitioners to evaluate evidence within and across intervention approaches. A review by Neely et al. (2018) evaluated the quality of the FCT generalization and maintenance literature, extending a review by Falcomata and Wacker (2013) focused on stimulus generalization. Studies were analyzed by their dimensions of generalization (i.e., generalization across tasks/activities, people, settings, or conditions), design in the generalization assessment (i.e., single probe post-treatment; multiple probes post-treatment

Table 1 Definitions of schedules of reinforcement terms

Terms	Definitions	Example
Delay-to-reinforcement	A procedure in which a child is taught to tolerate brief delays or denials to receive the requested reinforcer	A child requests an iPad. The iPad is given to the child after 5 s of her request. The 5 s is the brief delay
Chained schedule of reinforcement	A procedure in which the child is taught to complete a series of tasks prior to receiving a reinforcer	A child asks for a break. The child is provided a break only after she completes putting away her toys
Multiple schedules of reinforcement	A procedure in which the child is taught to differentiate between periods of the day when reinforcers may or may not be accessible using signals	A teacher uses a green card to signal to the child the toy is available. Once the child requests the toy, the teacher delivers the toy. The teacher uses a red card to signal to the child the toy is unavailable. If the child requests the toy, the teacher does not deliver the toy
Response restriction	A removal of the communication mode when the reinforcers cannot be delivered to the child	When group instruction begins, the teacher removes the child's picture card that says "toys please"
Alternative activity	During the times of the day in which reinforcers cannot be delivered to the child, the child is provided with an alternative activity	When the child requests bubbles during periods of time when it is unavailable, the teacher gives the child another activity (e.g., playdough)

Definitions adapted from Hagopian et al. (2011)

either with or without pre-treatment probe data; or continuous probes before, during, and after treatment), design in the maintenance assessment (i.e., single or multiple probe), generalization programming/teaching strategy (e.g., program common stimuli), maintenance probe latency, and results (e.g., positive, mixed, or negative outcomes on generalization). The researchers used the pilot standards for single-case research developed by the What Works Clearinghouse (WWC; Kratochwill et al. 2013) to assess study quality. The researchers developed their own rubric for evaluating the quality of the FCT generalization and maintenance literature based on the WWC standards, considering the unique practices that take place for probing and programming generalization and maintenance.

Neely et al. (2018) determined that, although many studies were sufficient according to the WWC standards, few studies met their adapted standards for generalization and maintenance, most of which did not use generalization and maintenance programming techniques used in the practice of applied behavior analysis (Stokes and Baer 1977). In other words, it was most common for research studies on FCT to “train and hope” for generalization (i.e., not program for generalization). Of the six articles that met the adapted WWC standards for generalization or met them with reservations, four demonstrated positive results, three of which included no generalization programming. The FCT generalization data was found to be lacking important quality indicators set for interrater reliability and a minimum number of data per phase.

Purpose Statement

There is no consensus in the field regarding quality standards for evaluating generalization and maintenance phases of single-case research studies; this may be why procedures and quality represented in the literature varies (Neely et al. 2018). Considering this, the current meta-analysis was conducted to be inclusive of FCT studies evaluating approaches to promote sustained behavior change. This allowed us to maximize the samples of data included within effect size analyses for variables of interest. This study was designed to extend Hagopian et al.’s (2011) review by updating the literature on FCT with schedule thinning techniques to analyze data descriptively, according to quality standards, and meta-analytically. We also sought to analyze follow-up phases of interest in studies according to their quality with procedures proposed by Neely et al. (2018), while not restricting meta-analysis inclusion by the level of quality within the maintenance/schedule thinning phase. Although there are limitations to meta-analyzing single-case data that do not meet quality standards, the purpose of this study was to determine preliminarily important directions for future

FCT schedule thinning literature. A final purpose of this review was to evaluate the literature within the early childhood population (birth to 8 years) of individuals with IDD. Findings were restricted to this population to make specific conclusions for practitioners working with children with IDD. Heath et al. (2015) found that FCT outcomes generally appeared to be most positive with young children in comparison to adults. However, we are extending Heath’s review by including moderators such as schedule thinning procedures, settings, interventionists, and quality level of the studies to determine factors that may impact the effectiveness of FCT for children with IDD.

This review aimed to meta-analyze single-case studies that examined thinning schedules of reinforcement following FCT in children with IDD ages 8 years and younger. The focus was restricted to children ages 8 and younger based on the definition of early childhood by the Council for Exceptional Children’s Division for Early Childhood (DEC). Additionally, we evaluated the rigor of the included studies based on Reichow’s (2011) quality indicators. Specifically, the research questions were: (a) What is the quality of studies that included a thinning schedule procedure following FCT based on Reichow’s (2011) indicators? (b) Which study variables moderated the effects of FCT followed by a thinning schedule of reinforcement on challenging behavior? and (c) Which study variables moderated the effects of FCT followed by a thinning schedule of reinforcement on FCRs?

Method

Search Procedure

The following three online databases were used to locate studies that incorporated a thinning reinforcement procedure following FCT: Google Scholar, ERIC, and PsycINFO. Multiple searches were completed by pairing keywords from the following two categories: (a) autism (search terms: ‘autism,’ ‘autism spectrum disorder,’ ‘disability,’) and (b) functional communication training (search terms: ‘communication training,’ ‘functional communication,’ ‘functional analysis communication,’ and ‘synthesized functional analysis’). In other words, we paired each search term from the first category with each search term from the second category using “and” as well as “or.” The search was limited to peer-reviewed journal articles published in English. We limited the search to peer-reviewed articles to improve the likelihood of including high-quality studies in the review and to be consistent with prior meta-analyses (e.g., Cowan et al. 2017; Ledbetter-Cho et al. 2018) published within *Journal of Autism and Developmental Disorders*. The reference lists of published literature reviews on FCT were also reviewed to identify potentially relevant studies (i.e., Andzik et al.

2016; Walker et al. 2018; Chezan et al. 2018; Durand and Moskowicz 2015; Falcomata and Wacker 2013; Gerow et al. 2018; Heath et al. 2015; Mancil 2006; Neely et al. 2018; Tiger et al. 2008).

In addition, a hand search was conducted to identify studies published between 1985 (when FCT was established by Carr and Durand) to 2018 in the *Journal of Applied Behavior Analysis*, a journal with a history of publishing behavior-analytic intervention studies. This journal was selected based on Smith (2012) review findings that showed the *Journal of Applied Behavior Analysis* to be the most common source of published studies employing single-case designs. Finally, the reference lists of all included studies were reviewed. The search process concluded in June of 2018. After duplicates were removed, 863 articles remained to be screened for eligibility. An initial title and abstract review to exclude extraneous articles (e.g., literature reviews, dissertations, books chapters) resulted in a total of 216 potentially-relevant articles. The fifth author conducted another independent search using the same search procedures for inter-rater agreement (IRA). IRA was calculated by dividing the number of agreements by the number of agreements and disagreements multiplied by 100. As the fifth author also located the same 28 studies, and did not find additional studies, the IRA was 100%.

Inclusion and Exclusion Criteria

Abstracts of the 216 articles were reviewed to determine whether an article met the inclusion criteria. When the abstract did not clearly state the use of a reinforcement thinning procedure following FCT, the authors accessed the full text to gather information about the independent variables and inspected the graphs to determine whether the study incorporated a reinforcement thinning procedure. Likewise, when the abstract did not state the behaviors targeted in the intervention, the authors read the dependent measure section of article to determine whether the article reported a challenging behavior measure. Qualifying studies met the following criteria: (a) included at least one participant between the ages of 2 and 8 years who was diagnosed with an intellectual disability or other developmental disability (e.g., ASD, Down syndrome), as coding was completed at the participant level, (b) used a reinforcement thinning procedure following FCT (i.e., multiple schedules, chained schedules, delay-to-reinforcement, alternative activities, response restriction), and (c) included a dependent measure for challenging behavior. Challenging behavior included any behavior that was identified as problematic by the authors such as aggression (e.g., hitting, kicking), self-injury (e.g., head banging, self-pinching), property destruction (e.g., throwing objects), disruption (e.g., crying, yelling), non-compliance, and/or elopement.

A study was excluded when it met at least one of the following exclusion criteria: (a) did not include at least one participant age 8 years or younger (e.g., Kahng et al. 1997), (b) did not include participants with a diagnosis of an intellectual or other developmental disability (e.g., Petscher and Bailey 2008), (c) did not include a reinforcement thinning procedure following FCT (e.g., Muharib et al. 2019), and/or (d) used additional strategies (e.g., noncontingent reinforcement, punishment). If a study used additional strategies for some but not all of the participants, the study was included but the participants who received such intervention were excluded (e.g., Jan in Fisher et al. 2000). Based on the inclusion and exclusion process, a total of 28 studies were included in the review. To collect IRA, the fifth author was randomly assigned with 30% of the 216 articles to determine whether a study met or did not meet the inclusion criteria. The IRA for the inclusion and exclusion process was 100%.

Data Extraction and Analysis

Data Coding

We coded and summarized the included studies in terms of (a) participant descriptions (i.e., age, diagnosis, communication level), (b) challenging behavior (i.e., aggression, self-injury, elopement, property destruction, disruption), (c) communication form selected for the participant during FCT [i.e., vocal, augmentative and alternative communication (AAC), or both], (d) functional behavior assessments [FBAs; i.e., functional analysis (FA), descriptive FBAs], (e) targeted functions for intervention (i.e., escape, attention, tangibles, automatic, access to rituals), (f) research design (i.e., reversal, multiple baseline/probe, alternating treatment, changing criterion), (g) setting (i.e., home, school, clinic), (h) interventionist (i.e., researcher, parent), and (i) dependent variables measured during the reinforcement thinning intervention. The first author extracted data across all studies. To collect IRA, the fifth author completed data extraction on a code-by-code basis across 30% of the studies selected at random. The IRA was 99% (range 97–100%) for data extraction. Disagreements were discussed and resolved by reviewing the variables in the articles.

Quality Appraisal

We used the quality indicators for single-case studies suggested by Reichow (2011) to determine the quality of the included studies. We used Reichow's indicators for their high rigor and detailed criteria for each indicator. As described by Reichow, primary quality indicators include (a) participant characteristics, (b) independent variable, (c) baseline condition, (d) dependent variable, (e) visual analysis, and (f) experimental control. Secondary quality indicators are

(a) interobserver agreement, (b) Kappa, (c) raters who were naïve to the purpose of the review, (d) fidelity, (e) generalization or maintenance, and (f) social validity. Each study was coded with high, acceptable, or unacceptable on each of the primary quality indicators and coded with a yes or no on each secondary quality indicators.

Based on the guidelines set forth by Reichow (2011), a single-case study can have a strong, adequate, or weak strength of quality. For a study to receive a strong quality rating, the study had to (a) meet all the primary quality indicators by receiving a high quality rating on each and (b) meet three or more secondary quality indicators. For a study to receive an adequate quality rating, the study had to (a) receive a high quality rating on at least four primary indicators, (b) receive no unacceptable quality rating on any of the primary indicators, and (c) meet at least two secondary quality indicators. A study received a weak quality rating when it (a) received a high quality rating on fewer than four primary indicators or (b) met fewer than two secondary indicators. Quality evaluations were completed by the first author. To calculate IRA, the fifth author evaluated 30% of the studies selected at random; these studies were different from those chosen for IRA during data extraction. IRA was calculated by adding the number of agreements (the final decision of weak, adequate, or strong), dividing the total by the number of reviewed articles, and then multiplying by 100. The IRA between the authors was 100%.

Preparation of Data

Although there is no agreed-upon method for calculating effect sizes in single-case research (Ledford et al. 2014), we used Tau-U for its ability to control for positive baseline trend. Tau-U (Parker et al. 2011) is a non-overlap method that was developed to address the issue of previous non-overlap methods. In addition to controlling for positive baseline trend, Tau-U can handle small data sets and discriminate magnitudes at the upper and lower limits (Vannest and Ninci 2015). In addition, studies included in this meta-analysis are not varied in their characteristics (e.g., communication level, communication form, age and setting). For instance, the majority of our cases received treatment in a clinical setting. Due to the small number of the included studies in each category and subcategory, procedures such as meta-regression are not a recommended option (Borenstein et al. 2009). Finally, the use of confidence interval is a common practice in single-case meta-analyses as demonstrated in a variety of publications (e.g., Chaffee et al. 2017; Chezan et al. 2018; Cumming and Rodríguez 2017; Tincani and De Mers 2016; Whalon et al. 2015).

To calculate Tau-U scores, we extracted the value of each data point in participant graphs by using UN-SCAN-IT version 5.2 (Silk 1992). This program allows one to manually

digitize underlying *x*- and *y*-axis data points when data point values are not reported. To extract data from the graphs representing withdrawal, changing criterion, multielement, multiple-baseline, or multiple-probe designs, the second author identified each adjacent AB pair (baseline and following intervention phase); each pair was treated separately. Similarly, when a combination of withdrawal and multiple-baseline or multiple-probe designs was present, data from each adjacent AB pair was extracted.

From the 28 studies included in this meta-analysis, a total of 270 AB phase contrasts were documented (extracted data can be requested from the first author). Fourteen studies used reversal designs or variations thereof (i.e., ABAB [$n=5$], ABABC [$n=4$], ABAC [$n=1$], ABCBC [$n=1$], ABAB-CDADADAD [$n=1$], ABABCACDAD [$n=1$], ABCDEFABCDEF [$n=1$]), four studies used changing criterion designs, five studies implemented multiple-baseline designs (i.e., across conditions [$n=1$], across therapists and settings [$n=1$], across two pairs and participants [$n=1$], across participants [$n=1$], and across rituals [$n=1$]), and five studies used a combination of two designs (i.e., reversal design with multielement [$n=3$], multiple baseline with reversal design [$n=2$]). In studies using reversal designs, AB contrasts were identified by pairing each (A) phase with the consecutive (B) phase. If a study included additional phases, such as a (C) phase or a (D) phase, the baseline condition (A) was paired with each included phase.

Effects Size Calculations

To calculate Tau-U, data were entered into the Tau-U calculator at www.singlecaseresearch.org. All baselines were corrected and both baseline and comparison phases were combined to obtain Tau-U scores. We used the “weighted” feature in the calculator to obtain a weighted average of all the previous Tau-U scores. We used the weighted mean rather than the mean of all phase contrast means in order to calculate the overall mean, as not all phase contrast means had the same “weight” (i.e., they had different number of data points). Tau-U scores range from -1 to 1 (Parker et al. 2011) and can be interpreted using the following criteria: (a) 0.20 or lower suggests a small effect; (b) between 0.20 and 0.60 suggests a moderate effect; (c) 0.60 to 0.80 suggests a large effect; and (d) above 0.80 suggests a very large effect (Vannest and Ninci 2015). IRA was collected on effect size calculations. The fifth author calculated the effects size for 30% of the contrasts. IRA results were 100%.

Moderator Analysis

After calculating Tau-U for each phase contrast, effect sizes were compared within each potential moderator. Two hundred seventy phase contrasts were used for this

meta-analysis. One hundred forty-two contrasts targeted challenging behavior reduction and 128 targeted FCR acquisition. Nine potential moderators were selected for analysis. Moderators pertaining to the characteristics of participants were: (a) grade level (preschool vs. elementary); (b) ages (2, 3, 4, 5, 6, 7, and 8); (c) comorbidity (single vs. multiple diagnoses), and (d) communication level (full sentences, defined as more than one-word-sentences; single words; gestures, defined as any prelinguistic behavior such as pointing or leading; and single words and gestures). Potential moderators pertaining to the characteristics of the intervention were (a) procedures (delay-to-reinforcement, chained schedules of reinforcement, multiple schedules of reinforcement, alternative activities, response restriction), (b) settings (home, school, clinic), (c) interventionists (researchers, parents), and (d) communication forms selected for the participant (vocal, AAC, or both). The final potential moderator pertained to the quality of the studies based on Reichow's (2011) indicators (adequate vs. weak). The effect of each potential moderator was separately analyzed for (a) challenging behavior and (b) FCRs.

The moderator analysis was completed in three steps. First, two Microsoft Excel spreadsheets were created for each moderator (e.g., setting, age). One of the spreadsheets included the phase contrasts pertaining to challenging behavior, and the other spreadsheet included the phase contrasts pertaining to FCRs. Next, we calculated the omnibus effect size for each category by adding all effect sizes within each spreadsheet. In order to evaluate the confidence interval ranges and determine statistical significance, the authors set the upper and lower bounds to 83.4%. At 83.4%, confidence interval ranges between variables represents $p < .05$ (Payton et al. 2003). Finally, all the omnibus effect sizes were placed in two tables, including the following values: (a) Tau-U, (b) upper confidence interval, and (c) lower confidence interval set at 83.4%. See Tables 3 and 4 for a summary of Tau-U values.

Results

Table 2 provides a descriptive summary of the studies. In the following section, we describe results from both descriptive and moderator analyses.

Descriptive Analysis

Participant Characteristics

A total of 51 participants between 2 and 8 years old received FCT across the 28 included studies. We coded the participants in terms of age, diagnosis, and communication level.

Age We coded the participant age group as preschool (ages 2 to 5) and elementary (ages 6 to 8). Over half of the participants ($n = 32$, 62.7%) were in the preschool age range, and 37.3% ($n = 19$) were in the elementary age group.

Diagnosis Slightly over half ($n = 26$, 50.9%) had a diagnosis of ASD with or without other secondary diagnoses. Thirty-seven percent of participants ($n = 19$) had a diagnosis of an intellectual or other developmental disability with or without a secondary diagnosis. A few participants ($n = 11$, 21.5%) were diagnosed with a behavioral disorder such as attention deficit hyperactivity disorder, oppositional defiant disorder, or disruptive disorder. The number do not add up to 51 as some participants had multiple diagnoses.

Communication Level Over half ($n = 29$, 56.8%) had been communicating using full sentences (more than one word), 11.7% ($n = 6$) using a single word, 9.8% ($n = 5$) using both single words and gestures, and 7.8% ($n = 4$) using gestures. A communication level was not reported for seven participants.

Intervention Characteristics

We coded nine variables in terms of intervention characteristics. These were topography of challenging behavior, FCR form selected for the participant, FBAs, targeted functions, intervention, dependent measures, research design, setting, and interventionist.

Challenging Behavior Aggression was the most common topography of challenging behavior exhibited by participants ($n = 40$, 78.4%). Forty-three percent ($n = 22$) of the participants engaged in disruptive behavior, 35.2% ($n = 18$) engaged in property destruction, and 29.4% ($n = 15$) engaged in self-injurious behavior. Only 7.8% ($n = 4$) displayed elopement. As some participants engaged in multiple topographies of challenging behavior, the numbers do not add up to 51 and the percentages do not add up to 100%.

FCR We coded the FCR forms selected for FCT intervention as either vocal, AAC, or both. For over half of the participants ($n = 30$, 58.8%), a vocal response was taught and for 23.5% ($n = 12$), an AAC response (a picture, speech generating device, or sign) was taught. Only 3.9% of the participants ($n = 2$) were taught to mand using two forms (vocal and AAC). It should be noted that, for some participants, an FCR form was not reported (see Table 2).

FBAs Experimental and/or descriptive FBAs were used to identify the function(s) of participants' challenging behavior. With the majority of participants ($n = 43$, 84.3%), FA as described by Iwata et al. (1994) was used to identify the

Table 2 Summaries of the included studies according to reinforcement thinning procedures

Article	Participants (name, age in years, diagnoses, communication level)	Problem behavior	FCR form	Functional behavior assessments	Target function (s)	Type of research design	Interventionist	Setting	Dependent variables	Study rigor
Delay to reinforcement										
Beaulieu et al. (2018)	Ally, 7, ASD, full sentences.	Agg, SIB, elopement, disruption	Vocal	Interview-informed synthesized FA	Singular; tangible	ATD with initial baseline	Researcher	Home	Problem behavior, FCR, tolerance response	Weak
Braithwaite and Richdale (2000)	Michael, 7, ASD, ID, full sentences..	Agg, SIB	Vocal	Interviews, observations	Multiple; tangible, attention	MB across 2 conditions	Researcher	School	Problem behavior, FCR	Weak
Fisher et al. (2000)	Ken, 3, cerebral palsy, ID, gestures	Property destruction	AAC	FA	Singular; attention	Reversal (ABAB)	Researcher	Clinic	Problem behavior, FCR	Weak
Ghaemmaghani et al. (2016)	Jack, 2, ASD, single words and gestures. Alex, 6, ASD, full sentences	Agg, disruption	Vocal and AAC, vocal	Interview-informed synthesized FA	Multiple; attention, escape, tangible	Reversal (ABAC)	Researcher	Clinic	Problem behavior, FCR, tolerance response	Weak
Hanley et al. (2014)	Gail, 3, PDD-NOS, full sentences. Bob, 8, ASD, full sentences	Agg, disruption	Vocal	Interview-informed synthesized FA	Multiple; attention, tangible (escape from redirection and access to tangible for Bob)	Changing criterion	Researcher	Clinic	Problem behavior, FCR, tolerance response	Adequate
Jessel et al. (2018)	Steve, 4, ASD, NR	Elopement	AAC	Interview-informed synthesized FA	Singular; tangible	Changing criterion	Researcher	Clinic	Problem behavior, FCR, tolerance response	Weak
Muehling et al. (2018)	Alex, 5, ASD, NR	Agg	AAC	FA	Multiple; attention, escape	Reversal (ABCBC)	Researcher	Home	Problem behavior, FCR	Weak
Rispoli et al. (2014)	Timmy, 4, PDD-NOS, full sentences. John, 3, ASD, NR. Diego, 3, PDD-NOS, full sentences	Agg, disruption, property destruction, SIB	Vocal, AAC.	FA (with pairwise comparisons)	Singular; access to rituals	MP across rituals	Researcher/parent (Timmy)	Clinic/home (Timmy)	Problem behavior, FCR	Adequate

Table 2 (continued)

Article	Participants (name, age in years, diagnoses, communication level)	Problem behavior	FCR form	Functional behavior assessments	Target function(s)	Type of research design	Interventionist	Setting	Dependent variables	Study rigor
Strand and Eldevik (2018)	John, 4, ASD, full sentences	Agg, disruption, SIB, elopement	Vocal	Interview-informed synthesized FA	Multiple; tangible, attention	Changing criterion	Researcher	Home	Problem behavior, FCR, tolerance response	Adequate
Chained schedules of reinforcement										
Berg et al. (2007)	Calvin, 5, DD, single words & gestures. Brent, 4, DD, full sentences. Theo, 4, DD, full sentences. Alex, 4, pervasive DD, speech and language disorder, single words and gestures	Agg, non-compliance, property destruction	Vocal, AAC	Scatter plots, A-B-C observations, and FA	Multiple; Escape, tangible, attention	Reversal (ABABC)	Parents	Home	Problem behavior, FCR, task completion, social interactions	Adequate
Davis et al. (2018)	Noah, 8, ADHD, ODD, and disruptive mood dysregulation disorder, full sentences	Agg	Vocal	Interviews, FA	Singular; escape	Reversal (ABABC)	Researcher	Clinic	Problem behavior, FCR, task completion	Weak
Falcomata et al. (2012a)	Steph, 8, Asperger's, full sentences. Mike, 8, ASD, full sentences	Agg, property destruction, SIB, disrupting, disruption	Vocal	FA	Singular; tangible	Reversal (ABAB)	Researcher	Clinic	Problem behavior, FCR	Weak
Falcomata et al. (2012b)	Danny, 8, ASD, full sentences.	Agg, SIB, property destruction	Vocal	FA	Multiple; escape, attention, tangible	Reversal (ABAB)	Researcher	School	Problem behavior, FCR	Weak
Falcomata et al. (2013)	Alonzo, 7, ASD, full sentences.	Agg, disruption	Vocal	FA	Multiple; escape, tangible	Reversal (ABAB)	Researcher	School	Problem behavior, FCR	Weak
Hagopian et al. (1998)	Case 17, 6, ID, NR.	Agg, SIB, disruption	NR	FA	Singular; tangible	Reversal (ABABC) with ATD	Researcher	Clinic	Problem behavior	Weak

Table 2 (continued)

Article	Participants (name, age in years, diagnoses, communication level)	Problem behavior	FCR form	Functional behavior assessments	Target function (s)	Type of research design	Interventionist	Setting	Dependent variables	Study rigor
Schieltz et al. (2011)	Juan, 4, ASD, DD, single words. Cam, 3, DD, single words. Bud, 3, ASD, ID, single words	Agg, SIB	Vocal	FA	Singular; escape	Reversal (ABABCDADADAD)	Parents	(Not indicated)	Problem behavior	Weak
Suess et al. (2014)	Lane, 3, PDD-NOS, full sentences. Jace, 2, PDD-NOS, ID, gestures. Jude, 3, PDD-NOS, full sentences	Agg, SIB, property destruction	Vocal, AAC	FA	Singular; escape	ATD	Parents	Home	Problem behavior	Weak
Wacker et al. (2011)	Tina, 4, ASD, DD, single words. Jose, 4, Fragile X, ID, single words	Agg, property destruction, SIB	Vocal	FA	Singular; escape	Reversal (ABABCAC-DAD)	Parents	Home	Problem behavior, FCR, task completion	Weak
Multiple schedules of reinforcement										
Betz et al. (2013)	Casey, 5, pervasive DD, full sentences Gordon, 6, ADHD, adjustment disorder, full sentences Henry, 6, OCD, full sentences	Agg, disruption, elopement	Vocal	FA	Singular; tangible	Reversal (ABAB)	Researcher	Clinic	Problem behavior, FCR	Weak
Fisher et al. (2015)	Corey, 5, PDD-NOS, full sentences Jacob, 6, ODD, full sentences	Agg, property destruction, elopement	Vocal	FA (with pairwise comparisons)	Singular; tangible	MB across therapists and clinic rooms	Researcher	Clinic	Problem behavior, FCR	Weak

Table 2 (continued)

Article	Participants (name, age in years, diagnoses, communication level)	Problem behavior	FCR form	Functional behavior assessments	Target function(s)	Type of research design	Interventionist	Setting	Dependent variables	Study rigor
Fuhrman et al. (2016)	Ben, 5, unspecified disruptive behavior, impulse control, conduct disorder, full sentences.	Agg, SIB, property destruction, pica	Vocal, AAC	FA (with pairwise comparisons)	Singular; tangible	Reversal (ABCDEFAB-CDEF)	Researcher	Clinic	Problem behavior, FCR	Weak
Greer et al. (2016)	Case 17, 6, ADHD, NR, Case 10, 4, ASD, NR	Agg, SIB, disruption	NR	FA	Singular; escape, or tangible (one per case)	Reversal (ABABC) (ABABCD)	Researcher	Clinic	Problem behavior, FCR	Weak
Kuhn et al. (2010)	Angela, 8, autistic disorder, stereotypic movement disorder, ID, single words and gestures	SIB, disruption	Vocal	FA (with pairwise comparisons)	Singular; attention	MB across two pairs and participants	Researcher	Clinic	Problem behavior, FCR	Weak
Shamlan et al. (2016)	Maurice 5, ASD, full sentences. Bernard, 5, ASD, full sentences	Disruption	Vocal	FA (with pairwise comparisons)	Singular; tangible	MB across participants	Researcher	Clinic	Problem behavior, FCR	Weak
Response restriction										
Falcomata et al. (2010)	John, 5, ASD, single words.	Elopement	Vocal	FA	Singular; automatic	MB across settings with an embedded reversal (ABAB)	Researcher	Clinic	Problem behavior, FCR, door play	Adequate

Table 2 (continued)

Article	Participants (name, age in years, diagnoses, communication level)	Problem behavior	FCR form	Functional behavior assessments	Target function (s)	Type of research design	Interventionist	Setting	Dependent variables	Study rigor
Fisher et al. (2014)	Evan, 7, ASD, stereotypic movement disorder, disruptive disorder, single words and gestures. Dan, 7, ID, disruptive behavior disorder, full sentences. John, 4, ASD, disruptive behavior disorder, full sentences. Casey, 4, ASD, intermittent explosive disorder, full sentences	Agg, SIB, property destruction, tantrums, elopement	AAC	FA	Singular; escape	MB across participants with an embedded reversal (ABCBCB)	Researcher	Clinic	Problem behavior, FCR	Weak
Greer et al. (2016)	Case 10, 4, ASD, NR. Case 6, 6, ASD, ID, NR	Agg, SIB, disruption	NR	FA	Singular; attention, or escape (one per case)	Reversal (ABABCD)	Researcher	Clinic	Problem behavior, FCR	Weak
Roane et al. (2004)	Juan, 7, ASD, ID, gestures	Agg	AAC	FA	Singular; attention	Reversal (ABABC)	Researcher	Clinic	Problem behavior, FCR	Weak
Hagopian et al. (2005)	Matt, 7, ASD, ID, ADHD, full sentences	Agg	Vocal and AAC	FA	Singular; attention	Reversal (ABABC) with ATD	Researcher	Clinic	Problem behavior, FCR	Weak

AAC augmentative and alternative communication, ADHD attention deficit hyperactivity disorder, ASD autism spectrum disorder, ATD alternating treatments design, Agg aggression, DD developmental disability, FA functional analysis, FCR functional communicative response, ID intellectual disability, MB multiple baseline, MP multiple probe, NR not reported, OCD obsessive compulsive disorder, ODD oppositional defiant disorder, PDD-NOS pervasive developmental disorder not otherwise specified, SIB self-injurious behavior, W weak, A adequate, and S Strong. Greer et al.'s (2016) study was included twice as it incorporated two procedures

Table 3 Analysis and outcomes for challenging behavior

Moderator	Tau-U	Upper CI	Lower CI
Grade level			
Preschool	-0.44	-0.490	-0.396
Elementary	-0.57	-0.636	-0.509
Age			
2	-0.45	-0.174	-0.727
3	-0.40	-0.471	-0.338
4	-0.64	-0.743	-0.530
5	-0.42	-0.514	-0.333
6	-0.47	-0.600	-0.331
7	-0.65	-0.735	-0.558
8	-0.52	-0.641	-0.397
Setting			
Clinic	-0.53	-0.573	-0.480
Home	-0.40	-0.486	-0.317
School	-0.98	-1.181	-0.779
Study rigor			
Adequate	-0.92	-1.014	-0.829
Weak	-0.40	-0.445	-0.362
Interventionist			
Parents	-0.31	-0.378	-0.237
Researcher	-0.96	-1.165	-0.759
Comorbidity			
No	-0.52	-0.574	-0.469
Yes	-0.45	-0.508	-0.398
Communication form			
AAC	-0.56	-0.620	-0.496
Vocal	-0.50	-0.546	-0.447
Vocal and AAC	-0.11	-0.055	-0.284
Communication level			
Gestures	-0.50	-0.558	-0.445
Full sentences	-0.56	-0.620	-0.496
Single words	-0.11	-0.055	-0.284
Gestures and single words	-0.48	-0.583	-0.376
Study category			
Chained schedule of reinforcement	-0.37	-0.434	-0.314
Delay of reinforcement	-0.69	-0.772	-0.616
FCT and response restriction	-0.57	-0.710	-0.428
FCT and alternative activity	-0.82	-1.202	-0.447
FCT and multiple schedule	-0.45	-0.519	-0.379

Tau-U small change: <0.20; moderate change: 0.20–0.60; large change: 0.60–0.80; very large change: >0.80

function(s) of participants' challenging behavior whether alone or subsequent to descriptive FBAs. For 9.8% of these participants ($n=5$), FAs were conducted in addition to interviews and/or observations. With 13.7% ($n=7$), interview-informed functional analysis as described by Hanley et al. (2014) was used. With one participant only (1.9%), descriptive FBAs were used and not followed by an FA.

Table 4 Analysis and Outcomes for FCR

Moderator	Tau-U	Upper CI	Lower CI
Grade level			
Preschool	0.11	0.056	0.165
Elementary	0.56	0.506	0.620
Age			
2	0.74	1.018	0.466
3	0.71	0.604	0.821
4	0.49	0.400	0.588
5	0.30	0.206	0.387
6	0.18	0.076	0.275
7	0.19	0.109	0.278
8	0.14	0.261	0.026
Setting			
Clinic	0.20	0.155	0.242
Home	0.63	0.523	0.730
School	0.58	0.379	0.779
Study rigor			
Adequate	0.73	0.643	0.813
Weak	0.15	0.107	0.195
Interventionist			
Parents	0.92	0.790	1.058
Researcher	0.14	0.038	0.249
Comorbidity			
No	0.40	0.344	0.452
Yes	0.14	0.079	0.193
Communication form			
AAC	0.20	0.137	0.273
Vocal	0.37	0.318	0.419
Vocal and AAC	0.25	0.406	0.092
Communication level			
Gestures	0.20	0.137	0.273
Full sentences	0.37	0.313	0.420
Single words	0.39	0.220	0.559
Gestures and single words	0.25	0.406	0.092
Intervention			
Chained schedule of reinforcement	0.47	0.371	0.576
Delay of reinforcement	0.44	0.371	0.516
FCT and response restriction	0.00	-0.119	0.110
FCT and alternative activity	0.61	0.989	0.231
FCT and multiple schedule	0.19	0.131	0.251

Tau-U small change: <0.20; moderate change: 0.20–0.60; large change: 0.60–0.80; very large change: >0.80

Targeted Functions For the majority of participants ($n=37$, 72.5%), challenging behavior was controlled by a singular form of reinforcement whereas for 27.5% of participants ($n=14$), challenging behavior was multiply controlled. FCT was delivered to 52.9% of participants ($n=27$) whose challenging behavior was maintained by access to tangibles, 47% of participants whose challeng-

ing behavior was maintained by escape from demands or attention ($n=24$), and 31.3% of participants whose challenging behavior was maintained by access to attention $n=16$). For only one participant (1.9%), challenging behavior was maintained by automatic reinforcement, and for 5.8% ($n=3$), challenging behavior was maintained by access to rituals. The percentages do not add up to 100% as challenging behavior of 14 participants served multiple functions.

Intervention Thirty-five percent of participants ($n=18$) received a chained schedule of reinforcement, 25.4% ($n=13$) received delay-to-reinforcement, 23.5% ($n=12$) received a multiple schedule of reinforcement, 1.9% ($n=1$) received an alternative activity procedure in addition to multiple schedules of reinforcement, and 15.6% ($n=8$) received a response restriction intervention following FCT. One participant received two types of intervention, therefore, the numbers do not add up to 51 and the percentages do not add up to 100%.

Dependent Measures For all participants ($n=51$, 100%), challenging behavior was the primary dependent measure. For 86.2% ($n=44$), FCRs were also measured during reinforcement thinning procedures. Less frequently, a tolerance response (e.g., “OK”) was measured for 13.7% of participants ($n=7$), and task completion was measured for 11.7% ($n=6$).

Research Design For over half the participants ($n=34$, 66.6%), a reversal design, whether alone or combined with another research design, was used to demonstrate the effects of a reinforcement thinning procedure. Researchers evaluated the effects of the interventions using a multiple baseline design for 27.4% of participants ($n=14$), alternating treatment design for 7.8% ($n=4$), whether alone or combined with other research designs, and changing criterion design for 5.9% ($n=3$) of participants. The percentages do not add up to 100% as for some participants, a combination of designs was used.

Setting and Interventionist Most participants ($n=32$, 62.7%) received the intervention in a clinical setting. For 25.5% of participants ($n=13$), the intervention was delivered in their homes. Only 5.9% of participants ($n=3$) received the intervention in a school setting. A setting was not reported for three participants. In the reviewed studies, researchers and parents served as interventionists. For 74.5% of participants ($n=38$), the intervention was delivered by a researcher whereas only 25.5% of participants ($n=13$) received the intervention by a parent.

Study Quality

Of the 28 studies, only 28.6% ($n=8$) demonstrated adequate quality based on Reichow’s (2011) criteria. The remaining studies ($n=20$, 71.4%) were of weak quality. Of particular interest, only 10.7% of the studies ($n=3$) reported data on procedural fidelity (Beaulieu et al. 2018; Rispoli et al. 2014; Suess et al. 2014). In terms of generalization, only 7.1% of studies ($n=2$) measured generalization of the skills (Beaulieu et al. 2018; Shamlan et al. 2016).

Moderator Analyses

Table 3 presents moderator effect sizes and confidence intervals for FCR. Table 4 presents moderator effect sizes and confidence intervals for challenging behavior.

Participant Characteristics

Moderator analyses of participant characteristics included age (i.e., 2, 3, 4, 5, 6, 7, 8), grade level (i.e., preschool, elementary), secondary diagnosis, communication level (i.e., full sentences, single words, gestures, gestures and single words) and communication form (i.e., AAC, vocal, vocal and AAC) categories. Ages ranged from 2 to 8 years and most participants were 4 years old (25%). Thirty (58.8%) participants were identified as being in preschool, whereas 21 (41.2%) participants were identified as being in elementary school. Effects by age categories ranged from small to large effect for FCR and moderate to large effect for challenging behavior. These effect sizes show a wider range for FCR (0.14 to 0.74) in comparison to challenging behavior (0.40 to 0.65). Results did not differ based on age category. However, the largest effects are shown in FCR for ages 2 and 3 whereas the lowest effect is shown in FCR at age 8. This pattern seems to indicate that FCR is more likely to be successful when participants are young. Effects by grade level ranged from small to moderate effect for FCR and moderate effect for challenging behavior. Like in age categories, FCR showed a wider range. Larger effects were seen in elementary compared to preschool.

Effects sizes for secondary diagnosis ranged from small to moderate for FCR and were moderate for challenging behavior. Participants ($n=19$) with secondary diagnoses showed a smaller effect ($ES=0.29$) than participants ($n=32$) who did not have an additional diagnosis ($ES=0.46$). Overall, outcomes for communication level show a larger effect for FCR ($ES=0.82$) than for challenging behaviors ($ES=0.60$). Interestingly, single words produced the largest effect when the goal was to decrease challenging behaviors ($ES=0.39$) but produced the smallest effect when the goal was to increase FCR ($ES=0.11$). Effect sizes for communication form ranged from small to moderate. For both FCR and

challenging behaviors, vocal and AAC had the lowest effect sizes ($ES=0.25$ and 0.11). AAC had the highest effect size for challenging behaviors ($ES=0.56$) whereas vocal had the highest effect size for FCR ($ES=0.37$).

Intervention Characteristics

Moderator analyses of intervention characteristics included intervention type, settings (i.e., clinic, home, school), and interventionist (i.e., parents, researchers) categories. Effects for intervention type ranged from small effect to large effect for FCR and moderate to very large effect for challenging behavior. Of all the interventions, alternative activities produced the largest effect sizes for both FCR and challenging behavior ($ES=0.61$ and 0.82), though it should be emphasized that this was based on only one participant in the sample. Chained schedules of reinforcement produced the lowest effect size for challenging behavior ($ES=0.37$) and multiple schedules of reinforcement produced the lowest effect size for FCR ($ES=0.19$). Because response restriction entailed removal of the FCR mode, the effect size was 0 for FCR. Effects for setting ranged from small to large effect for FCR and moderate to large effect for challenging behavior. FCR showed again a wider range. These results show an interesting pattern. When it comes to decreasing challenging behaviors, school settings seem to show much larger effect ($ES=0.98$) and home settings show the lowest effect ($ES=0.40$). When the focus is to increase the FCR, home settings seem to show larger effect ($ES=0.63$). In this case, school setting shows a lower effect but not far behind ($ES=0.58$). These outcomes seem to indicate that home settings are better suited for the acquisition of skills. Effects for interventionist ranged from small effect to very large effect for FCR and moderate to very large effect for challenging behavior. Outcomes also show an interesting pattern. When the goal of studies was to increase FCRs, parents were more effective ($ES=0.92$) and researchers were less effective ($ES=0.14$). When the goal was to reduce challenging behaviors, the opposite was true. In this case, researchers were more effective ($ES=0.96$) and parents were less effective ($ES=0.31$). These outcomes seem to indicate that parents were more effective when teaching FCRs while researchers were more effective when the goal was to reduce challenging behaviors.

Study Rigor

Moderator analysis of study rigor included adequate and weak. Effects ranged from small to large for FCR and moderate to very large for challenging behavior. These outcomes show that studies with an adequate level of rigor show a larger effect size ($ES=0.82$) compared to studies with weak

rigor ($ES=0.27$). According to these results, studies with higher degree of quality produce larger effect sizes.

Discussion

The purpose of this meta-analytic review was to summarize studies in which thinning schedules of reinforcement following FCT were evaluated for children with IDD ages 8 and younger. Given the prevalence of challenging behavior among children with IDD (Emerson et al. 2001; Lowe et al. 2007) and the well-established evidence-base supporting FCT as strategy to address challenging behavior (e.g., Muharib and Wood 2018; Wong et al. 2013), it is important to explore the conditions under which thinning schedules of reinforcement following FCT has been implemented and whether such strategies are more effective in addressing challenging behavior when implemented under particular conditions, while also evaluating the quality of the supportive research. This information can inform guidelines for stakeholders who assume responsibility for implementing FCT with children in a range of settings (e.g., parents, teachers, therapists) and identify critical areas for future research. In the following sections, we describe key findings, implications, and future research areas.

Key Findings and Implications

We conducted descriptive analyses to summarize participant and study characteristics and study quality and moderator analyses to examine whether particular variables contributed to more or less pronounced child outcomes during reinforcement thinning conditions. Findings and implications for both sets of analyses are described in detail in the following sections.

Descriptive Outcomes

Overall, reinforcement thinning procedures were implemented across a wide range of participant and intervention conditions. However, a majority of children were reported to have advanced communication skills, with over half communicating vocally in full sentences. Children with severe IDD often have complex communication needs that necessitate AAC to replace or supplement speech and may communicate at less advanced levels (e.g., Snell et al. 2010). This is important to note, as children in the reviewed studies may not represent those considered to have severe disabilities; therefore implications related to our findings may be more applicable to those with less extensive support needs. We were unable to code for extent of support needs as information to do so (e.g., adaptive behavior scale results, IQ assessment scores) was typically not reported.

A majority of interventions were informed by an experimental FA with only one informed by descriptive FBA. Typically, experimental FA is considered the most efficient and precise strategy for identifying behavioral function(s) and informing subsequent intervention planning (Falcomata et al. 2012a). This finding is promising, as it increases the likelihood that the FCT and reinforcement thinning strategies were technically adequate and more likely to yield desired outcomes. However, it is important to consider implications for those implementing FCT in natural environments, as experimental FA is usually conducted by a highly-skilled assessor, and therefore may not be feasible or contextually appropriate for school and home settings. Interestingly, Walker et al. (2018) found that FCT involving AAC was more effective in school settings when informed by descriptive FBA as compared to FA, raising important questions about the utility of different FBA strategies in natural settings. Nonetheless, when faced with complex cases, FA should be considered, with skilled experts providing training or assistance when appropriate (e.g., Rispoli et al. 2015; Simacek et al. 2017).

Findings also revealed that a variety of reinforcement thinning procedures identified in previous reviews (Hagopian et al. 2011; Tiger et al. 2008) have been explored with young children with IDD. Caregivers and teachers should consider a range of reinforcement thinning strategies based on contextual factors and student characteristics (Tiger et al. 2008) but should exercise caution when selecting strategies that limit opportunities to communicate (even if temporarily). For example, response restriction involves the removal of a child's mode of communication when reinforcement is unavailable, a strategy that when applied to students who use AAC can significantly interfere with their communication rights (see Brady et al. 2016). In most cases, reinforcement thinning strategies were used to address aggression, a finding similar to that of Hagopian et al. (2011), with fewer cases focused on disruptive behavior, property destruction, self-injury, and elopement. Young children engage in a wide range of behaviors that can interfere with learning and social interactions, damage property, and/or cause harm to the child or others (Powell et al. 2007). If unresolved, challenging behavior can increase in intensity and occurrence, and may become more resistant to intervention as the child ages (Heath et al. 2015). As such, it is important for practitioners to consider FCT followed by reinforcement thinning for any challenging behavior including, but not limited to aggression, that significantly interferes with daily functioning.

Another noteworthy finding relates to intervention setting. A majority of interventions was implemented in clinical settings, with fewer implemented in home and school environments. Specifically, only 13 and three children received FCT in home and school settings, respectively. This presents important questions about the extent to which child

outcomes generalize from clinical to natural settings, including home, school, and community environments. It is crucial for interventionists to program for generalization (Stokes and Baer 1977), as children with IDD often fail to generalize newly acquired skills across different conditions when “train and hope” approaches are employed. Regardless of setting, highly-skilled researchers primarily delivered the intervention. Therefore, implications of this review are significantly limited for school- and home-based intervention. In a few cases, family members were trained to implement FCT and reinforcement thinning procedures to acceptable levels of implementation fidelity (e.g., Suess et al. 2014), suggesting that individuals with limited or no experience in FCT can be successful in their implementation with external support involving performance feedback and prompting strategies. In fact, both Andzik et al. (2016) and Walker et al. (2018) found that teachers can effectively implement FCT across a range of students and challenging behaviors, though the extent to which this is true for reinforcement thinning following FCT is unknown.

Based on the quality appraisal following guidelines set forth by Reichow (2011), over a quarter of the studies demonstrated adequate quality, with the remaining studies demonstrating weak quality. Given this substantial number of studies with weak quality, readers should exercise caution when interpreting the results of the moderator analyses, as the credibility of the reinforcement thinning procedures outcomes may be comprised by weak quality ratings (Cooper et al. 2009). Furthermore, few studies measured the extent to which reinforcement thinning procedures were implemented with fidelity by interventionists, a common limitation across FCT studies (Muharib and Wood 2018). Because implementation fidelity can affect intervention effectiveness (Mayer et al. 2014), interventionists must carefully monitor the extent to which reinforcement thinning procedures are being implemented and adjust accordingly; otherwise, interventionists may erroneously attribute limited improvement in student behavior to an ineffective intervention plan, potentially leading to premature abandonment of the intervention.

Moderator Outcomes

Across all children, the reinforcement thinning procedures following FCT produced an overall moderate effect (0.49 for challenging behavior and 0.56 for FCRs). We examined effect sizes across several variables to determine whether they moderated reinforcement thinning intervention effect. In this section, we highlight a few notable findings from these analyses. Effect sizes for challenging behavior and FCRs were higher for students with a single diagnosis as compared to students with a secondary diagnosis. Perhaps students with a secondary diagnosis have more extensive support needs that affect communication abilities due to the

co-occurring disability, potentially limiting FCR skill acquisition and challenging behavior, though findings related to communication characteristics as moderators of FCT outcomes in other reviews are mixed (e.g., Walker et al. 2018; Heath et al. 2015).

Alternative activities as a reinforcement thinning approach produced the largest overall effect size for measures of challenging behavior and FCRs, with effect sizes for chained and multiple schedules of reinforcement the lowest for challenging behavior and FCRs, respectively. The alternative activities approach involves providing the child with an alternative activity (e.g., a toy) when the functional reinforcer is not available resulting in attenuation of motivation to engage in challenging behavior. Chained schedules of reinforcement involve demands whereby the child completes a series of activities before gaining access to the reinforcer, which can potentially be aversive depending on the nature of the demands. Multiple schedules of reinforcement often include a contingency specifying rule (e.g., wearing a specific colored wristband to signal reinforcement is available/unavailable and explaining the contingency to the child), and depending on a child's receptive language abilities, the contingency may not be well understood. To improve the effectiveness of these two strategies, practitioners can assess the child's support needs to make necessary adjustments (e.g., visual within-activity schedule of required activities prior to reinforcement). It should be noted that only one student received an alternative activity across the reviewed studies, thereby the extent to which conclusions can be drawn from this particular analysis is severely limited.

In addition, effect sizes for reinforcement thinning were highest for challenging behavior measures when researchers delivered the intervention, whereas effect sizes were highest for FCRs when parents implemented the intervention. This is an interesting outcome as parents are considered natural communication partners and may have more influence over communication skill development due to their histories of interaction with their children (Biggs and Meadan 2018), whereas researchers are unfamiliar communication partners without such histories but are highly skilled in behavioral intervention and are more likely to implement procedures with high levels of fidelity. Similarly, effect sizes were the highest for FCRs when implemented in the home setting, likely due to the fact that family members tended to implement interventions in home environments. Reductions in challenging behavior were greater in school settings compared to other settings, possibly due to exposure to peer models or additional supports in place (e.g., school- and class-wide behavior strategies). As was the case earlier, the limited number of cases in which reinforcement thinning following FCT was implemented in school and home settings limits the extent to which one can draw conclusions from this analysis.

A final outcome from the moderator analyses revealed that effect sizes typically were stronger for studies with higher quality ratings. One particular area of study quality mentioned earlier was the absence of implementation fidelity measurement across studies. If the extent to which an intervention is being implemented with fidelity is assessed over the course of an intervention and adjustments are made when implementation drops below an acceptable level, it is more likely that desirable student outcomes will be achieved. Furthermore, in order to a practice to be identified as evidence-based, it must be implemented with fidelity and measurement of such must be present in research demonstrations (Horner et al. 2005).

Limitations and Future Research Directions

There are a few limitations that are important to consider that inform future research directions. We conducted a moderator analysis by descriptively comparing Tau-U effect size scores across potential moderator variables. As such, we were unable to detect statistically significant differences in effect sizes, a process that could strengthen the findings. However, our primary purpose was to preliminarily evaluate potential moderators so as to identify additional areas for research that will lead to a more robust literature base. In terms of the search procedures, we did not search the gray literature for qualifying dissertations and master thesis papers, which could lead to publication bias, as there has been limited consensus on the acceptability of including unpublished research in meta-analytic reviews as demonstrated in recent meta-analyses that included the gray literature (e.g., Maggin et al. 2017; Mason et al. 2013) and those that did not (e.g., Cowan et al. 2017; Ledbetter-Cho et al. 2018). However, with the increasing acceptance of and encouragement to include gray literature in systematic reviews (e.g., Gage et al. 2017), it might be helpful to include these additional demonstrations of reinforcement thinning evaluations to provide a more comprehensive review of the work in this area.

There are several significant limitations in the literature itself that will need to be explored further. For example, the limited number of cases in which reinforcement thinning was implemented school settings warrants additional research to explore the feasibility and effectiveness of different thinning strategies within natural settings where other children and adults are present. Similarly, the literature contains few examples of typical interventionists (e.g., teachers, family members). Research efforts should focus on implementation of FCT thinning procedures among these interventionists, while also exploring the type and dosage of training necessary to produce desired outcomes. There is a growing body of work supporting the effectiveness of training practices such as coaching with performance feedback (Fallon et al. 2015), with such work extending to FCT implementation in

school settings (e.g., Andzik et al. 2016; Walker et al. 2018). There was only one instance in which alternative activities was used to thin reinforcement post FCT. More research is needed to explore this promising practice. Finally, we found it difficult to code for certain child participant characteristics, as information was unclear or unavailable. In particular, the extent of children's support needs was unclear; thus, it was difficult to determine for whom reinforcement thinning procedures were most effective.

Conclusion

We reviewed and meta-analyzed 28 intervention studies that involved a thinning schedule procedure following FCT for children with IDD ages 8 and younger. The results of Tau-U analyses demonstrated overall moderate effect sizes for both challenging behavior and FCRs. The findings suggested that thinning procedures were most effective for children who had stronger communication repertoires. Although the results suggested strong effects of alternative activities, additional research is needed as the sample consisted of only one child who had received such treatment. Overall, the findings from this review are promising and provide preliminary guidance for practitioner implementation of thinning schedule procedures following FCT.

Author Contributions RM contributed to the formulation of the review questions and procedures, conducted literature searches, provided descriptive summaries of the included studies, and contributed to the writing of the manuscript. FA contributed to the formulation of the review procedures, data coding and analysis and, the writing of the manuscript. JN contributed to the formulation of the review procedures and writing of the manuscript. VLW contributed to the writing of the manuscript. APV contributed to literature searches and data coding and analysis. All authors have approved the final version of the manuscript.

Compliance with Ethical Standards

Conflicts of interest The authors declare that they have no conflicts of interest.

References

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

- Andzik, N. R., Cannella-Malone, H. I., & Sigafoos, J. (2016). Practitioner-implemented functional communication training: A review of the literature. *Research and Practice for Persons with Severe Disabilities, 41*, 79–89. <https://doi.org/10.1177/1540796916633874>.
- *Beaulieu, L., van Nostrand, M. E., Williams, A. L., & Herscovitch, B. (2018). Incorporating interview-informed functional analyses into practice. *Behavior Analysis in Practice, 11*(4), 385–389.
- Berg, W., Wacker, D., Harding, J., Ganzer, J., & Barretto, A. (2007). An evaluation of multiple dependent variables across distinct classes of antecedent stimuli pre and post functional communication training. *Journal of Early and Intensive Behavioral Intervention, 3*, 305–333.
- Betz, A. M., Fisher, W. W., Roane, H. S., Mintz, J. C., & Owen, T. M. (2013). A component analysis of schedule thinning during functional communication training. *Journal of Applied Behavior Analysis, 46*, 219–241.
- Biggs, E. E., & Meadan, H. (2018). Early communication interventions for young children with intellectual and developmental disabilities: The roles of natural communication partners. *International Review of Research in Developmental Disabilities, 55*, 1–37. <https://doi.org/10.1016/bs.iridd.2018.08.005>.
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester: Wiley.
- Brady, N. C., Bruce, S., Goldman, A., Erickson, K., Mineo, B., Ogle-tree, B. T., ... Wilkinson, K. (2016). Communication services and supports for individuals with severe disabilities: Guidance for assessment and intervention. *American Journal on Intellectual and Developmental Disabilities, 121*, 121–138.
- *Braithwaite, K. L., & Richdale, A. L. (2000). Functional communication training to replace challenging behaviors across two behavioral outcomes. *Behavioral Interventions: Theory & Practice in Residential & Community-Based Clinical Programs, 15*(1), 21–36.
- Carr, E. G., & Durand, V. M. (1985). Reducing behavior problems through functional communication training. *Journal of Applied Behavior Analysis, 18*(2), 111–126.
- Chaffee, R. K., Briesch, A. M., Johnston, A. H., & Volpe, R. J. (2017). A meta-analysis of class-wide interventions for supporting student behavior. *School Psychology Review, 46*, 149–164. <https://doi.org/10.17105/SPR-2017-0015.V46-2>.
- Chezan, L. C., Wolfe, K., & Drasgow, E. (2018). A meta-analysis of functional communication training effects on problem behavior and alternative communicative responses. *Focus on Autism and Other Developmental Disabilities, 33*, 195–205. <https://doi.org/10.1177/1088357617741294>.
- Cooper, H., Hedges, L., & Valentine, J. (2009). *The handbook of research synthesis and meta-analysis*. New York: Russell Sage Foundation.
- Cowan, R. J., Abel, L., & Candel, L. (2017). A meta-analysis of single-subject research on behavioral momentum to enhance success in students with autism. *Journal of Autism and Developmental Disorders, 47*, 1464–1477. <https://doi.org/10.1007/s10803-017-3076-6>.
- Cumming, T. M., & Rodríguez, C. D. (2017). A meta-analysis of mobile technology supporting individuals with disabilities. *The Journal of Special Education, 51*, 164–176. <https://doi.org/10.1177/0022466917713983>.
- Davis, T. N., Weston, R., Hodges, A., Uptegrove, L., Williams, K., & Schieltz, K. M. (2018). Functional communication training and demand fading using concurrent schedules of reinforcement. *Journal of Behavioral Education, 27*, 343–357.
- Durand, V. M., & Moskowitz, L. (2015). Functional communication training: Thirty years of treating challenging behavior. *Topics in Early Childhood Special Education, 35*, 116–126. <https://doi.org/10.1177/0271121415569509>.
- Emerson, E., Kiernan, C., Alborz, A., Reeves, D., Mason, H., Swarbrick, R., ... Hatton, C. (2001). The prevalence of challenging behaviors: A total population study. *Research in Developmental Disabilities, 22*, 77–93.

- Falcomata, T. S., Muething, C. S., Gainey, S., Hoffman, K., & Fragale, C. (2013). Further evaluations of functional communication training and chained schedules of reinforcement to treat multiple functions of challenging behavior. *Behavior Modification*, *37*, 723–746. <https://doi.org/10.1177/0145445513500785>.
- Falcomata, T. S., Roane, H. S., Feeney, B. J., & Stephenson, K. M. (2010). Assessment and treatment of elopement maintained by access to stereotypy. *Journal of Applied Behavior Analysis*, *43*, 513–517. <https://doi.org/10.1901/jaba.2010.43-513>.
- *Falcomata, T. S., Roane, H. S., Muething, C. S., Stephenson, K. M., & Ing, A. D. (2012a). Functional communication training and chained schedules of reinforcement to treat challenging behavior maintained by terminations of activity interruptions. *Behavior Modification*, *36*(5), 630–649.
- Falcomata, T. S., White, P., Muething, C. S., & Fragale, C. (2012b). Functional communication training and a chained schedule of reinforcement procedure to treat challenging behavior with multiple functions. *Journal of Developmental and Physical Disabilities*, *24*, 529–538.
- Falcomata, T. S., & Wacker, D. (2013). On the use of strategies for programming generalization during functional communication training: A review of the literature. *Journal of Developmental and Physical Disabilities*, *25*, 5–15. <https://doi.org/10.1007/s10882-012-9311-3>.
- Fallon, L. M., Collier-Meek, M. A., Maggin, D. M., Sanetti, L. M. H., & Johnson, A. H. (2015). Is performance feedback for educators an evidence-based practice? A systematic review and evaluation based on single-case research. *Exceptional Children*, *81*, 227–246. <https://doi.org/10.1177/0014402914551738>.
- Fisher, W. W., Greer, B. D., Fuhrman, A. M., & Querim, A. C. (2015). Using multiple schedules during functional communication training to promote rapid transfer of treatment effects. *Journal of Applied Behavior Analysis*, *48*, 713–733.
- Fisher, W. W., Greer, B. D., Querim, A. C., & DeRosa, N. (2014). Decreasing excessive functional communication responses while treating destructive behavior using response restriction. *Research in Developmental Disabilities*, *35*, 2614–2623. <https://doi.org/10.1016/j.ridd.2014.06.024>.
- *Fisher, W. W., Thompson, R. H., Hagopian, L. P., Bowman, L. G., & Krug, A. (2000). Facilitating tolerance of delayed reinforcement during functional communication training. *Behavior Modification*, *24*(1), 3–29.
- Fuhrman, A. M., Fisher, W. W., & Greer, B. D. (2016). A preliminary investigation on improving functional communication training by mitigating resurgence of destructive behavior. *Journal of Applied Behavior Analysis*, *49*, 884–899.
- Gage, N. A., Cook, B. G., & Reichow, B. (2017). Publication bias in special education meta-analyses. *Exceptional Children*, *83*, 428–445. <https://doi.org/10.1177/0014402917691016>.
- Gerow, S., Hagan-Burke, S., Rispoli, M., Gregori, E., Mason, R., & Ninci, J. (2018). A systematic review of parent-implemented functional communication training for children with ASD. *Behavior Modification*, *42*, 335–363. <https://doi.org/10.1177/0145445517740872>.
- *Greer, B. D., Fisher, W. W., Saini, V., Owen, T. M., & Jones, J. K. (2016). Functional communication training during reinforcement schedule thinning: An analysis of 25 applications. *Journal of Applied Behavior Analysis*, *49*(1), 105–121.
- Hagopian, L. P., Boelter, E. W., & Jarmolowicz, D. P. (2011). Reinforcement schedule thinning following functional communication training: Review and recommendations. *Behavior Analysis in Practice*, *4*, 4–16. <https://doi.org/10.1007/BF03391770>.
- Hagopian, L. P., Fisher, W. W., Sullivan, M. T., Acquisto, J., & LeBlanc, L. A. (1998). Effectiveness of functional communication training with and without extinction and punishment: A summary of 21 inpatient cases. *Journal of Applied Behavior Analysis*, *31*, 211–235. <https://doi.org/10.1901/jaba.1998.31-211>.
- *Hagopian, L. P., Kuhn, S. A. C., Long, E. S., & Rush, K. S. (2005). Schedule thinning following communication training: Using competing stimuli to enhance tolerance to decrements in reinforcer density. *Journal of Applied Behavior Analysis*, *38*(2), 177–193.
- *Hanley, G. P., Jin, C. S., Vanselow, N. R., & Hanratty, L. A. (2014). Producing meaningful improvements in problem behavior of children with autism via synthesized analyses and treatments. *Journal of Applied Behavior Analysis*, *47*(1), 16–36.
- Heath, A. K., Ganz, J. B., Parker, R., Burke, M., & Ninci, J. (2015). A meta-analytic review of functional communication training across mode of communication, age, and disability. *Review Journal of Autism and Developmental Disorders*, *2*, 155–166. <https://doi.org/10.1007/s40489-014-0044-3>.
- Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children*, *71*, 165–179.
- Iwata, B. A., Dorsey, M. F., Slifer, K. J., Bauman, K. E., & Richman, G. S. (1994). Toward a functional analysis of self-injury. *Journal of Applied Behavior Analysis*, *27*, 197–209. (Reprinted from *Analysis and Intervention in Developmental Disabilities*, *1982*, *2*, 3–20).
- Jessel, J., Ingvarsson, E. T., Metras, R., Whipple, R., Kirk, H., & Solsbery, L. (2018). Treatment of elopement following a latency-based interview-informed, synthesized contingency analysis. *Behavioral Interventions*, *33*, 271–283. <https://doi.org/10.1002/bin.1525>.
- Kahng, S., Iwata, B. A., DeLeon, I. G., & Worsdell, A. S. (1997). Evaluation of the “control over reinforcement” component in functional communication training. *Journal of Applied Behavior Analysis*, *30*(2), 267–277.
- Kuhn, D. E., Chirighin, A. E., & Zelenka, K. (2010). Discriminated functional communication: A procedural extension of functional communication training. *Journal of Applied Behavior Analysis*, *43*, 249–264. <https://doi.org/10.1901/jaba.2010.43-249>.
- Kratochwill, T. R., Hitchcock, J. H., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., et al. (2013). Single-case intervention research design standards. *Remedial and Special Education*, *34*, 26–38. <https://doi.org/10.1177/0741932512452794>.
- Ledbetter-Cho, K., O’Reilly, M., Lang, R., Watkins, L., & Lim, N. (2018). Meta-analysis of tablet-mediated interventions for teaching academic skills to individuals with autism. *Journal of Autism and Developmental Disorders*, *48*, 3021–3036. <https://doi.org/10.1007/s10803-018-3573-2>.
- Ledford, J. R., Wolery, M., & Gast, D. L. (2014). Controversial and critical issues in single case research. In D. L. Gast & J. R. Ledford (Eds.), *Single case research methodology: Applications in special education and behavioral sciences* (pp. 377–396). New York: Routledge.
- Lowe, K., Allen, D., Jones, E., Brophy, S., Moore, K., & James, W. (2007). Challenging behaviours: Prevalence and topographies. *Journal of Intellectual Disability Research*, *51*(8), 625–636.
- Maggin, D. M., Pustejovsky, J. E., & Johnson, A. H. (2017). A meta-analysis of school-based group contingency interventions for students with challenging behavior: An update. *Remedial and Special Education*, *38*, 353–370. <https://doi.org/10.1177/0741932517716900>.
- Mancil, G. R. (2006). Functional communication training: A review of the literature related to children with autism. *Education and Training in Developmental Disabilities*, *41*, 213–224.
- Mason, R. A., Davis, H. S., Boles, M. B., & Goodwyn, F. (2013). Efficacy of point-of view video modeling: A meta-analysis. *Remedial and Special Education*, *34*, 333–345. <https://doi.org/10.1177/0741932513486298>.

- Mayer, G. R., Sulzer-Azaroff, B., & Wallace, M. (2014). *Behavior analysis for lasting change* (3rd ed.). Cornwall-on-Hudson, NY: Sloan Publishing.
- Muehling, C. S., Falcomata, T. S., Ferguson, R., Swinnea, S., & Shpall, C. (2018). An evaluation of delay to reinforcement and mand variability following functional communication training. *Applied Behavior Analysis, 51*, 263–275.
- Muharib, R., Correa, V. I., Wood, C. L., & Haughney, K. L. (2019). Effects of functional communication training using GoTalk Now iPad application on challenging behavior of children with autism spectrum disorder. *Journal of Special Education Technology, 34*, 71–79. <https://doi.org/10.1177/0162643418783479>.
- Muharib, R., & Wood, C. L. (2018). Evaluation of the empirical support of functional communication training for children with autism spectrum disorder. *Review Journal of Autism and Developmental Disorders, 5*, 360–369. <https://doi.org/10.1007/s40489-018-0145-5>.
- Neely, L., Garcia, E., Bankston, B., & Green, A. (2018). Generalization and maintenance of functional communication training for individuals with developmental disabilities: A systematic and quality review. *Research in Developmental Disabilities, 78*. <https://doi.org/10.1016/j.ridd.2018.02.002>.
- Park, C., Yelland, G., Taffe, J., & Gray, K. (2012). Brief report: the relationship between language skills, adaptive behavior, and emotional and behavior problems in preschoolers with autism. *Journal of Autism and Developmental Disorders, 42*, 2761–2766. <https://doi.org/10.1007/s10803-012-1534-8>.
- Parker, R. I., Vannest, K. J., Davis, J. L., & Sauber, S. B. (2011). Combining nonoverlap and trend for single case research: Tau-U. *Behavior Therapy, 42*, 284–299. <https://doi.org/10.1016/j.beth.2010.08.006>.
- Payton, M. E., Greenstone, H. M., & Schenker, N. (2003). Overlapping confidence intervals or standard error intervals: What do they mean in terms of statistical significance? *Journal of Insect Science, 3*, 234. <https://doi.org/10.1093/jis/3.1.34>.
- Petscher, E. S., & Bailey, J. S. (2008). Comparing main and collateral effects of extinction and differential reinforcement of alternative behavior. *Behavior Modification, 32*, 468–488. <https://doi.org/10.1177/0145445507309032>.
- Powell, D., Fixsen, D., Dunlap, G., Smith, B., & Fox, L. (2007). A synthesis of knowledge relevant to pathways of service delivery for young children with or at risk of challenging behavior. *Journal of Early Intervention, 29*, 81–106.
- Reichow, B. (2011). Development, procedures, and application of the evaluative method for determining evidence-based practices in autism. In B. Reichow, et al. (Eds.), *Evidence-based practices and treatments for children with autism* (pp. 25–39). Boston: Springer.
- *Rispoli, M., Camargo, S., Machalicek, W., Lang, R., & Sigafos, J. (2014). Functional communication training in the treatment of problem behavior maintained by access to rituals. *Journal of Applied Behavior Analysis, 47*, 580–593.
- Rispoli, M., Ninci, J., Burke, M. D., Zaini, S., Hatton, H., & Sanchez, L. (2015). Evaluating the accuracy of results for teacher implemented trial-based function analyses. *Behavior Modification, 39*, 627–653. <https://doi.org/10.1177/0145445515590456>.
- *Roane, H. S., Fisher, W. W., Sgro, G. M., Falcomata, T. S., & Pabico, R. R. (2004). An alternative method of thinning reinforcer delivery during differential reinforcement. *Journal of Applied Behavior Analysis, 37*, 213–218.
- Schieltz, K. M., Wacker, D. P., Harding, J. W., Berg, W. K., Lee, J. F., Padilla Dalmau, Y. C., et al. (2011). Indirect effects of functional communication training on non-targeted disruptive behavior. *Journal of Behavioral Education, 20*, 15–32. <https://doi.org/10.1007/s10864-011-9119-8>.
- *Shamlian, K. D., Fisher, W. W., Steege, M. W., Cavanaugh, B. M., Samour, K., & Querim, A. C. (2016). Evaluation of multiple schedules with naturally occurring and therapist-arranged discriminative stimuli following functional communication training. *Journal of Applied Behavior Analysis, 49*, 228–250. Silk Scientific [UNSCANIT]. Utah: Orem
- Simacek, J., Dimian, A. F., & McComas, J. J. (2017). Communication intervention for young children with severe neurodevelopmental disabilities via telehealth. *Journal of Autism and Developmental Disorders, 47*, 744–767. <https://doi.org/10.1007/s10803-016-3006-z>.
- Smith, J. D. (2012). Single-case experimental designs: A systematic review of published research and current standards. *Psychological Methods, 17*(4), 510.
- Snell, M. E., Brady, N., McLean, L., Ogletree, B. T., Siegel, E., Sylvester, L., ... Sevcik, R. (2010). Twenty years of communication intervention research for individuals who have severe intellectual and developmental disabilities. *American Journal on Intellectual and Developmental Disabilities, 115*, 364–380. <https://doi.org/10.1352/1944-7558-115-5.364>.
- Stokes, T. F., & Baer, D. M. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis, 10*, 349–367.
- Strand, R. C., & Eldevik, S. (2018). Improvements in problem behavior in a child with autism spectrum diagnosis through synthesized analysis and treatment: A replication in an EIBI home program. *Behavioral Interventions, 33*, 102–111.
- *Suess, A. N., Romani, P. W., Wacker, D. P., Dyson, S. M., Kuhle, J. L., Lee, J. F., ... Waldron, D. B. (2014). Evaluating the treatment fidelity of parents who conduct in-home functional communication training with coaching via telehealth. *Journal of Behavioral Education, 23*, 34–59.
- Tiger, J. H., Hanley, G. P., & Bruzek, J. B. (2008). Functional communication training: A review and practical guide. *Behavior Analysis in Practice, 1*, 16–23.
- Tincani, M., & De Mers, M. (2016). Meta-analysis of single-case research design studies on instructional pacing. *Behavior Modification, 40*, 799–824. <https://doi.org/10.1177/0145445516643488>.
- Vannest, K. J., & Ninci, J. (2015). Evaluating intervention effects in single-case research designs. *Journal of Counseling & Development, 93*, 403–411. <https://doi.org/10.1002/jcad.12038>.
- *Wacker, D. P., Harding, J. W., Berg, W. K., Lee, J. F., Schieltz, K. M., Padilla, Y. C., ... Shahan, T. A. (2011). An evaluation of persistence of treatment effects during long-term treatment of destructive behavior. *Journal of the Experimental Analysis of Behavior, 96*, 261–282.
- Walker, V. L., Lyon, K. J., Loman, S. L., & Sennott, S. (2018). A systematic review of Functional Communication Training (FCT) interventions involving augmentative and alternative communication in school settings. *Augmentative and Alternative Communication, 34*, 118–129. <https://doi.org/10.1080/07434618.2018.1461240>.
- 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. <https://doi.org/10.1007/s10803-015-2373-1>.
- Wong, C., Odom, S. L., Hume, K. Cox, A. W., Fettig, A., Kucharczyk, S., ... Schultz, T. R. (2013). *Evidence-based practices for children, youth, and young adults with Autism Spectrum Disorder*. Chapel Hill: The University of North Carolina, Frank Porter Graham Child Development Institute, Autism Evidence-Based Practice Review Group.

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