The Effect of Parent-Implemented Functional Communication Training on Challenging Behavior and Communication: A Meta-Analysis

A DISSERTATION
SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

Brittany Pennington

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Dr. Jason Wolff, Adviser
Dr. Kristen McMaster, Adviser

May 2019

Acknowledgments

A dissertation takes a village, and this work would not be possible without the contributions of many people. First, thank you to my family- my parents, Mike and Kim Overstreet, who facilitated a lifelong love of learning and a desire to help make the world a better place. Thank you to my husband, Kyle, whose support and encouragement is unfailing, and to my daughter, Maggie, who brings me unparalleled joy.

Thank you to the mentors who, early in my career, encouraged me to pursue research and behavior analysis- Dr. Paul Hewitt, Dr. Sheila Barnes, and Dr. Elizabeth Lorah.

Thank you to my friends- in particular, thank you Dr. Jessica Simacek. I couldn't have done this without you and your work on parent-implemented FCT truly inspires me- PerSist! Thank you Jaclyn Gunderson for the many hours spent coding.

Thank you to my mentors in NRCLII- Drs. Blair Lloyd and Erin Barton- I learned a tremendous amount about research through my collaborations with you. Thank you to my committee, Drs. Jason Wolff, Kristen McMaster, Michael Rodriguez, Lizbeth Finestack, and Dan Maggin, for your guidance in the development and execution of this project. In particular, thank you to my advisors, Drs. Jason Wolff and Kristen McMaster. Thank you for consistently modeling exemplary research, for holding me to high standards, and for providing me with support and encouragement.

Dedication

This dissertation is dedicated to Kyle. You encouraged me to go to grad school (the first time, and the second time, and the third time…), brought me coffee when I was writing in the morning and dinner when I was writing in the evening, believed in me when I did not believe in myself, and distracted me when I needed distractions. I could not have finished this without you!

Abstract

A high proportion of people with developmental disabilities engage in challenging behavior compared to the general population (McClintock, Hall, & Oliver, 2003). Functional Communication Training (FCT) is an is an evidence-based intervention to address challenging behavior (Heath et al., 2015), but many people remain unable to access effective interventions like FCT. One strategy to increase access is to train parents to be interventionists for their children. The present study is a meta-analysis of studies examining parent-implemented FCT. Procedures were registered with PROSPERO prior to data extraction (Registration # CRD42018100912, Pennington, 2018). The study addresses the following questions: (1) What is the overall effect of parent-implemented FCT on challenging behavior and communication? (2) What are the characteristics of participants, implementers, and interventions in parent-implemented FCT studies, and to what extent do those characteristics moderate outcomes? (3) Do parents implement FCT with fidelity, and how were parents trained or coached? and (4) To what extent do included studies meet quality indicators? I used a multi-level, mixed effects meta-analysis to examine the effects of parent-implemented FCT on challenging behavior for 53 participants in 21studies, and on communication for 29 participants in 14 studies. Overall, FCT had a moderate to large effect size for reducing challenging behavior and a large effect size for increasing communication. No significant moderators were found for participant or coaching characteristics. For intervention characteristics, implementing the intervention in natural settings was significantly associated with an increased effect size. These results indicate that parent-implemented FCT is an effective intervention across various participant, intervention, and coaching characteristics.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Developmental Disabilities and Challenging Behavior

A high proportion of people with developmental disabilities engage in challenging behavior compared to the general population (McClintock, Hall, & Oliver, 2003). Challenging behavior includes behaviors that have the potential to cause injury or that require additional resources to manage, like aggression, self-injury, and property destruction (Emerson, 1995). Challenging behavior reduces quality of life (Murphy, 2009), limits community access (Worcester, Nesman, Mendez, & Keller, 2008), and stresses caregivers and family members (Lecavalier & Wiltz, 2016; Worcester et al., 2008).

According to Koritsas and Jacono's (2012) review of prevalence of and risk factors for challenging behavior, the most accurate estimate of the percentage of people with intellectual disabilities (ID) who engage in challenging behavior is 15-17.5%. This estimate is based on the British Prevalence study, which was conducted by Qureshi and Aborz in 1992 and replicated by Emerson and colleagues in 2001. In the initial study, Qureshi and Alborz (1992) selected seven representative districts in the UK, and identified day and residential facilities, schools, and healthcare facilities that served people with ID in those districts. Additionally, fieldwork teams identified others who had been excluded from those services. From each setting, researchers obtained a list of people with ID who were then screened for more information about challenging behavior. The researchers identified a sample that ranged from age 5 to 88, with a majority (60%)

between the ages of 12 and 35. They found that 16.7% of identified people engaged in challenging behavior, and among those who engaged in challenging behavior, 42% engaged in severe challenging behavior that had either caused injuries or property destruction, occurred at least weekly, or disrupted more than a few minutes at least once per day (Qureshi & Alborz, 1992). Emerson and colleagues (2001) repeated these procedures in two of the original districts, and found that 12% of identified people engaged in challenging behavior. Among those who engaged in challenging behavior, 64% engaged in severe aggression, and 33% engaged in severe self-injury (Emerson et al., 2001).

Challenging behavior is also prevalent among people with autism spectrum disorder (ASD). Kanne and Mazurek (2011) analyzed symptoms of ASD and aggression among 1,380 children between the ages of 4 and 17 years across multiple university sites as part of the Simons Simplex Collection. They measured aggression towards caregivers and towards others on a four-point scale that included: 0, no aggression; 1, mild aggression (i.e., threatening, rough play, or provoked lashing out); 2, aggression involving hitting or biting; and 3, violence including the use of implements. To clearly differentiate between definite aggressive behavior and no aggressive behavior, researchers split subjects into two groups: those who engaged in no aggression, indicated by a zero for both caregivers and others, and those who engaged in definite aggression, indicated by a two or three for caregivers and/or others. Fifty-six percent of their sample engaged in aggression at the time of the study, and 68% had engaged in aggression at some point in time (Kanne & Mazurek, 2011). Self-injurious behavior (SIB) is also

common among individuals with ASD. Duerden and colleagues (2012) evaluated SIB among 250 children and adolescents enrolled in genetic studies at two sites in Ontario, Canada. The sample included 212 boys and 38 girls who ranged in age from 21 months to 19 years. They assessed the presence of self-injury using relevant questions on the Autism Diagnostic Interview-Revised (ADI-R) and the Repetitive Behavior Scale-Revised (RBS-R), and found that 52% of participants engaged in some form of self-injury. Dimian and colleagues (2017) analyzed SIB among 235 toddlers at high familial risk for ASD recruited from across the United States. Participants' parents completed the RBS-R at 12 and 24 months. Approximately 39% of participants engaged in SIB at age 12 months, and 32% engaged in SIB at age 24 months.

Challenging Behavior and Communication

Given the high prevalence of challenging behavior among individuals with developmental disabilities, it is critical to examine how challenging behavior relates to other characteristics in order to inform intervention. A number of studies have established a relation between challenging behavior and communication deficits (e.g., Beitchman et al., 1996; Carson, Williams, Klee, Perry, & Lee, 1998). Beitchman and colleagues (1996) examined the relations between speech/language proficiency and behavior problems for 169 children. One-third of 5-year-old, English speaking kindergarteners in a district were randomly selected to participate in two stages of speech/language testing. Students who failed both stages of testing and a control sample of students who passed one or both tests received a third stage of developmental, behavioral, and psychiatric assessments. Of the 169 children who participated in all

stages of testing, 138 participated in follow-up testing at age 12. Children who had language impairments at age five demonstrated significantly greater behavior problems at age 12, indicating that early language delay might play a role in the emergence of later behavior problems.

Similarly, Carson and colleagues (1998) measured language proficiency, behavior problems, and social and cognitive development at ages two and three for 36 typically developing children. They recruited potential participants from a university database, and among potential participants, they selected a sample with a continuous distribution of language development scores. Scores of language proficiency at age two negatively predicted behavior problems at age three, accounting for 21% of the variance in the total challenging behavior score. Additionally, scores of language proficiency at age three negatively predicted behavior problems at age three, accounting for 35% of the variance in the total challenging behavior score. These results indicated a relation between deficits in communication at ages two and three and challenging behavior at age three.

Communication Intervention

Because of the relation between communication and challenging behavior, a promising strategy for intervening with students who engage in challenging behavior involves teaching communication (Carr & Durand, 1985; Walker & Snell, 2013). Walker and Snell (2013) conducted a meta-analysis of single-case research that measured the effectiveness of communication interventions on reducing challenging behavior. The researchers located 54 studies on augmentative and alternative communication (AAC) interventions, or interventions that involve teaching alternative communication

modalities like sign language, picture card exchange, or speech generating devices to nonverbal participants. They measured effect size using nonoverlap of all pairs (NAP), and found a mean effect size of .88 (*SD* = .18; range .11-1.0), indicating a moderate level of effect (Parker & Vannest, 2009). The researchers disaggregated results by participant characteristics, intervention characteristics, and interventionist characteristics, and found moderate to strong effects across all characteristics. Additionally, they conducted moderator analyses and found that participant age, type of pre-intervention assessment, and intervention type moderated outcomes. AAC interventions reduced challenging behavior more effectively for younger participants than older ones, interventions based on functional assessments reduced challenging behavior more effectively than interventions based on other assessments, and functional communication training (FCT) reduced challenging behavior more effectively than other interventions (Walker & Snell, 2013).

Functional Communication Training

FCT is a behavior intervention that is based on the idea that challenging behavior is functionally communicative, or that children engage in challenging behavior in order to influence their environment and meet their needs (Carr & Durand, 1985). Thus, challenging behavior can be reduced by teaching appropriate communication to serve the challenging behavior's function (Carr & Durand, 1985). FCT involves four parts: (1) identifying the function of challenging behavior, or the consequence that challenging behavior produces for the individual, (2) teaching a communication response to serve that function, (3) providing functional reinforcement following the communication response,

and (4) withholding functional reinforcement following challenging behavior (Tiger, Hanley, & Bruzek, 2008).

Carr and Durand (1985) conducted the seminal study on FCT. In Phase 1, the researchers identified contexts in which challenging behavior was likely to occur by presenting children with easy and difficult tasks, and varying the amount of adult attention across the easy tasks. This process enabled the researchers to compare challenging behavior across easy and difficult tasks with attention held constant, and across high and low amounts of attention with task difficulty held constant. Two students engaged in challenging behavior when presented with difficult tasks, one student engaged in challenging behavior when provided a low amount of attention, and one student engaged in challenging behavior in both the difficult-task and the low-attention conditions. In Phase 2, researchers used the conditions identified in Phase 1 to test the effects of FCT on challenging behavior. The researchers taught the participants to say, "I don't understand" and "Am I doing good work?" They provided help with the tasks when students said, "I don't understand," and praise when students said, "Am I doing good work?" For participants who engaged in challenging behavior when presented with difficult tasks, "I don't understand" was considered the relevant response and "Am I doing good work?" was considered the irrelevant response. For participants who engaged in challenging behavior when provided a low amount of attention, "Am I doing good work?" was considered the relevant response and "I don't understand" was considered the irrelevant response. The researchers alternated between baseline, relevant response, and irrelevant response conditions. All four participants engaged in the less

challenging behavior in the relevant response condition than in the irrelevant response condition or baseline. These results indicated that teaching a communicative response reduced challenging behavior in the context where challenging behavior was more likely to occur for all participants. Since Carr and Durand (1985) introduced FCT, many studies have confirmed the effectiveness of FCT for reducing challenging behavior, establishing FCT as an evidence-based practice (National Autism Center, 2015).

Barriers to Service Delivery

FCT is an established and effective evidence-based practice (Heath et al., 2015), but many people with disabilities who engage in challenging behavior remain unable to access effective interventions like FCT. Betz and colleagues (2012) analyzed the services accessed and the perceived barriers to services for families of 102 children with developmental disabilities. Only 8.8% of the sample accessed behavior management services, and 85.2% of the sample reported unaddressed behavior management concerns. The primary reasons participants listed for not accessing behavior management services were lack of information and a lack of available services (Betz et al., 2012).

Parents may find that needed services are difficult to access due to a shortage of providers. Wise and colleagues (2010) contacted all United States state and territory early intervention coordinators, and 52 (91%) participated in a survey about the demand for ASD services. Forty-six percent reported challenges in meeting the demand for initial service plans, and most reported shortages of personnel including behavior therapists (89%), speech-language pathologists (82%), and occupational therapists (79%). Among 34 respondents from states that reported the average number of service hours, nearly half

reported that children received five or fewer hours per week (Wise et al., 2010), far fewer than the 25-40 hours per week reported in EIBI research (Reichow, 2011).

In addition to affecting the intensity of available services, the provider shortage creates lengthy waitlists in many states. Dimian (2017) analyzed access to services among 667 young children diagnosed with ASD between 2008-2010 in Minnesota, and found an average of nine months between diagnosis and treatment onset. Yingling, Hock, and Bell (2017) analyzed access to services among 473 Medicaid recipients in South Carolina, and found an average of three years between diagnosis and treatment onset, with no differences across neighborhood or race-ethnicity. These delays to intervention are common, and many states report waitlists for services ranging from six months to several years (L&M Policy Research, 2014).

Cultural and geographic barriers may exacerbate difficulties accessing treatment. Hewitt and colleagues (2012) conducted in-depth, structured interviews about early intervention services for children with disabilities with 13 parents, 12 advocates, seven state or county staff, five clinicians, four residential service providers, four educators, and two attorneys in Minnesota. Stakeholders reported a lack of services among certain cultural groups and in rural areas. Several stakeholders reported persistent delays to diagnosis and intervention services in the Somali community, and one stakeholder reported that "ABA is nonexistent in greater Minnesota." Although Yingling et al. (2017) found no differences in delay to treatment onset based on race or neighborhood in South Carolina, Wise and colleagues (2010) used multivariate models to analyze provider shortages across states and found that states with a higher African American and

Hispanic population and states with a lower population density were able to offer fewer service hours than other states. Thus, despite the existence of effective behavior interventions like FCT, many families of children with disabilities may find services difficult to access due to a shortage of qualified personnel, lengthy waitlists, or cultural or geographic barriers.

Waiting for services adversely affects outcomes. Dimian (2017) cross-referenced early intensive behavior intervention (EIBI) service data and educational records. Experiencing a delay to EIBI significantly decreased the likelihood of a general education placement and of participating in the Minnesota Comprehensive Assessment, and significantly increased the odds of receiving an educational diagnosis of ASD and the number of special education service hours. Vivanti and Dissanayake (2016) found that children who started an EIBI program, the Early Start Denver Model, at 24 months achieved significantly larger verbal gains than students who started the same program at 48 months. Piccininni, Bisnaire, and Penner (2017) compared the cost of services for children with ASD given the current wait for services in Ontario (approximately 2.5 years), reduced wait, and no wait. The researchers concluded that eliminating the wait for services would save over $60,000 per year compared to the current wait.

Parent-Implemented Interventions

One strategy to increase access to services is to train parents to be interventionists for their children. A number of studies have demonstrated that researchers can coach parents to implement behavior interventions (e.g., Seuss et al., 2014; Wacker et al., 2013a; Wacker et al., 2013b). Wacker and colleagues (2013a, 2013b) coached parents to

implement FCT using telehealth. Researchers at the University of Iowa connected virtually via Skype with parents at regional clinics across the state. First, they coached parents to conduct functional analyses (FAs) to identify the function of challenging behavior for 20 participants between the ages of two and seven who had been identified with either pervasive developmental disorder-not otherwise specified (PDD-NOS) or autistic disorder (Wacker et al., 2013a). Researchers identified a function for 18 of the 20 participants, indicating that parents were able to conduct assessments that evaluated the function of challenging behavior. Then, researchers coached 13 of those parents and four additional parents through implementing FCT (Wacker et al., 2013b). All participants engaged in less challenging behavior in the final three FCT sessions compared to baseline, and on average, participants engaged in 93.5% less problem behavior during the final three FCT sessions than they did during baseline.

Seuss and colleagues (2014) evaluated the fidelity with which parents implemented FCT in their homes for three 2- to 3-year-old children diagnosed with PDD-NOS who engaged in self-injury, aggression, and property destruction. Behavior consultants virtually coached parents to conduct FCT, and parents made videos of trials in which they conducted the procedures independently without coaching. Researchers measured parents' fidelity to FCT procedures, including errors of commission and errors of omission. All three parents implemented FCT during independent trials with acceptable levels of fidelity (range 78%-87%), and two of the three participants engaged in little to no challenging behavior during the independent trials. Together, these findings indicate that parents can conduct FCT with acceptable levels of fidelity (Seuss et al.,

2014), and that parent-implemented FCT can produce decreases in challenging behavior (Seuss et al., 2014; Wacker et al., 2013a; Wacker et al., 2013b).

Despite a number of studies examining parent-implemented FCT, no meta-analysis has examined results of parent-implemented FCT across studies. Additionally, no meta-analysis or review article has looked across parent-implemented FCT studies to evaluate characteristics of parents, implementers, and interventions that could moderate outcomes. Meta-analyses provide stronger evidence that a practice is evidence-based than individual studies (Merlin, Weston, & Tooher, 2009). Given the need to identify evidence-based practices that can overcome barriers to service delivery and the promise of parent-implemented FCT, a meta-analysis on parent-implemented FCT is warranted.

Study Quality

In addition to examining effect sizes across studies, to generate recommendations and inform future research, it is necessary to evaluate the quality of parent-implemented FCT studies. The Council for Exceptional Children (CEC) produced Standards for Evidence-Based Practice in Special Education (2014). These quality indicators build on previous indicators proposed by Horner and colleagues (2005) and the What Works Clearinghouse (Kratochwill et al., 2013) and are comprehensive because they address multiple dimensions of quality. They address (a) whether each study provides sufficient detail for replication, (b) whether the independent variable is implemented with fidelity, (c) whether the design provides sufficient evidence of internal validity, (d) whether the dependent variable is appropriately selected and measured, and (e) whether data are appropriately analyzed. Additionally, Reichow, Barton, and Maggin (2015) developed

the Risk of Bias Assessment for Single-Case Design to evaluate the quality of single-case studies. These quality indicators address the potential that researcher bias affects the results of single-case research, including decisions about when to change conditions and whether participants and coders know when conditions change. These quality indicators reflect aspects of quality unique to single-case research, which tends to involve active researcher participation. Along with providing information about the body of research, quality indicators are useful in a meta-analysis because quality can be examined as a moderator to evaluate the extent to which study quality impacts effectiveness. To date, there are no published meta-analyses applying quality indicators to parent-implemented FCT studies.

The Present Study

The present study is a meta-analysis of studies examining parent-implemented FCT. The goal is to aggregate the power of individual studies, characterize the state of knowledge, and generate recommendations for future research. The present study addresses the following research questions:

1. What is the overall effect of parent-implemented FCT on challenging behavior and communication for children who engage in challenging behavior?
2. What are the characteristics of participants, implementers, and interventions in parent-implemented FCT studies?
3. To what extent do participant, intervention, or implementer characteristics moderate intervention effectiveness?

4. To what extent do parents implement FCT with fidelity?
5. How were parents trained or coached on implementation?
6. To what extent do fidelity or training/coaching characteristics relate to intervention effectiveness?
7. To what extent do included studies meet CEC and Risk of Bias quality indicators?

Chapter 2

Literature Review

In this review, I examined literature on FCT in order to define terms, identify components of FCT, and examine relevant previous review articles. Specifically, I defined FCT by examining the underlying behavioral principles to provide theoretical background and clarify terms. Next, I identified components of FCT that previous researchers have indicated may relate to outcomes. Then, I examined review articles on FCT to situate the present meta-analysis in the context of current literature. Finally, I examined review articles on parent training to identify training components potentially relevant to parent-implemented FCT.

Defining FCT

Behavioral principles. FCT is based on the principles of applied behavior analysis, which is the science of analyzing behavior by examining relations among behavior and environmental variables (Cooper, Heron, & Heward, 2007, p. 20). The goal of applied behavior analysis is to use the scientific study of behavior to inform behavior change programs that reduce challenging behavior and increase appropriate behavior (Baer, Wolf, & Risley, 1968). Below, I discuss the following terms and behavioral principles as they relate to FCT: behavior, consequence, antecedent, functional behavior assessment, function, differential reinforcement, and schedules of reinforcement.

Behavior analysts define *behavior* as an organism's observable, measurable movement that alters the environment (Johnston & Pennypacker, 1980). Behavior analysts examine overt behaviors rather than static states or private behaviors such as

thoughts or feelings. In the behavior analytic account, many behaviors are shaped over time by the consequences that they produce (Skinner, 1953, p. 65). FCT addresses consequence-shaped behaviors.

Behavior analysts examine the *consequences* that immediately follow behavior (Cooper et al., 2007, pp. 28 & 34). Several categories of consequences influence behavior: reinforcement, punishment, and extinction (Cooper et al., 2007, p. 37). *Reinforcement* includes any consequence that increases the likelihood that the behavior will occur in the future, such as obtaining a preferred item or escaping a non-preferred activity (Cooper et al., 2007, p. 36). *Punishment* includes any consequence that decreases the likelihood that the behavior will occur in the future, such as losing attention or a preferred item (Cooper et al., 2007, pp. 37-38). *Extinction* occurs when a reinforcer that has followed behavior previously is withheld (Cooper et al., 2007, p. 37). When an organism encounters extinction following behavior that previously produced reinforcement, the behavior may initially increase but then it becomes less and less frequent if extinction continues (Lerman & Iwata, 1996). FCT typically involves reinforcing communication and placing problem behavior on extinction in order to increase the frequency of communication and decrease the frequency of challenging behavior (Tiger et al., 2008).

In addition to consequences, behavior analysts examine the *antecedents* that precede behavior (Cooper et al., 2007, p. 28). Two categories of antecedents influence behavior: discriminative stimuli (SDs) and motivating operations (Cooper et al., 2007, pp. 37 & 375). *SDs* immediately precede behavior and signal the availability of

reinforcement (Cooper et al., 2007, p. 37). For example, if a child receives snacks from her father following challenging behavior when her father is in the kitchen, then father in the kitchen could serve as an SD signaling that challenging behavior may produce a snack. *Motivating operations* alter the value of a particular reinforcer (Cooper et al., 2007, p. 375). For example, a child may be more likely to engage in challenging behavior that produces snacks if he missed breakfast than if he had a large breakfast. In FCT, practitioners may attend to antecedent stimuli by using SDs to signal that communication will produce reinforcement (e.g., Landa & Hanley, 2016) or by conducting training when motivating operations make the reinforcer particularly valuable (Davis, Fuentes, & Durand, 2013).

Functional behavior assessment (FBA) is an umbrella term for a behavior analytic approach to assessing behavior that involves analyzing the antecedents and consequences that precede and follow the target behavior (Cooper et al., 2007, Chapter 24). Based on the idea that challenging behavior occurs because it receives reinforcement, FBAs are used to identify what reinforcement the challenging behavior produces. The reinforcement that a particular behavior produces is called the *function* of that behavior. Interviews or questionnaires such as the *Functional Assessment Interview* (FAI; O'Neill, Albin, Storey, Horner, & Sprague, 2015) involve caregivers answering questions about environmental events that may be related to the target behavior. *Descriptive assessment*s involve systematically observing and recording antecedents, behaviors, and consequences (e.g.. Lalli, Browder, Mace, & Brown, 1993). *Structured descriptive assessments* (SDAs) involve presenting antecedent conditions potentially

related to behavior in order to measure whether the target behavior occurs following the hypothesized antecedent (Anderson & Long, 2002). *Functional analysis* (FA), which is considered the gold standard approach to FBA, involves experimentally manipulating both antecedents and consequences (Iwata, Dorsey, Slifer, Bauman, & Richman, 1982/1994). In an FA, the researcher or interventionist alternates among sessions where she presents antecedents and provides reinforcement for the target behavior to determine whether behavior increases in the presence of hypothesized antecedents when it is followed by hypothesized consequences. Interventionists conduct an FBA prior to implementing FCT (Carr & Durand, 1985).

Differential reinforcement is a well-established, empirically supported intervention based on the results of an FBA (Petscher, Rey, & Bailey, 2009). After using an FBA to identify the function of challenging behavior, differential reinforcement involves: (1) providing the functional consequence following either the occurrence of a replacement behavior or the nonoccurrence of the challenging behavior, and (2) withholding the functional consequence following challenging behavior (Cooper et al., 2007, Chapter 22). FCT is a specific type of differential reinforcement in which the replacement behavior is a communication response (Carr & Durand, 1985).

Schedules of reinforcement refer to the specific conditions under which behaviors produce reinforcement (Cooper et al., 2007, p. 305). Schedules of reinforcement can be either fixed or varied and based on either intervals of time (interval schedules) or number of responses (ratio schedules; Cooper et al., 2007, p. 305). In a *fixed ratio* (FR) schedule, reinforcement follows a given number of occurrences of the target behavior, and in a

variable ratio (VR) schedule, reinforcement follows variable numbers of responses that average to a given mean (Cooper et al., 2007, pp. 306-307). In a *fixed interval* (FI) schedule, reinforcement follows the first occurrence of target behavior after a given interval of time, and in a *variable interval* (VI) schedule, reinforcement follows the first occurrence of a target behavior after variable intervals of time that average to a given mean (Cooper et al., 2007, pp. 310-312). Different schedules generally produce different rates of responding. FR schedules produce high rates of responding with pauses following reinforcement. VR schedules produce steady, high rates of responding. FI schedules produce responding that begins slowly then accelerates towards the end of the interval. VI schedules produce low to moderate, stable rates of responding (e.g., De Luca & Holborn, 1990/1992; Metzger & Lattal, 1998). Additionally, different schedules produce different patterns when behavior encounters extinction. Behavior maintained by variable schedules tends to persist longer in extinction than behavior maintained by fixed schedules, and behavior maintained by less-frequent reinforcement tends to persist longer in extinction than behavior maintained by frequent reinforcement (e.g., Doughty & Lattal, 2010; Kerns, 1975; Lattal, Reilly, Kohn, 2013). FCT typically involves an FR 1 schedule, indicating reinforcement is delivered after every one response, but some researchers transition to less frequent or variable schedules to increase persistence of the communication response (Tiger et al., 2008).

Components of FCT. Tiger and colleagues (2008) analyzed 91 articles on FCT with 204 participants in order to develop a practical guide for implementing FCT. Based on their review, Tiger and colleagues (2008) identified the following components of

FCT: assessing the function of the challenging behavior, selecting an appropriate communication response, selecting a context, teaching the communication response, responding to problem behavior, and reinforcing the communication response. Below, I describe each of the components described by Tiger and colleagues' (2008) review, and the recommendations they generated related to each component.

Practitioners implementing FCT must decide how to assess the function of the challenging behavior (Tiger et al., 2008). Researchers have evaluated the function using a variety of assessments, but the majority used functional analysis (Tiger et al., 2008). While assessments other than functional analysis may suffice in some situations, Tiger and colleagues (2008) recommended conducting a functional analysis prior to FCT for individuals who engage in severe problem behavior.

In addition to selecting a functional assessment methodology, interventionists must select a communication response. Tiger and colleagues (2008) outlined three considerations related to response selection. First, practitioners should consider the effort required to emit the communication response. Initially, the communication response should be less effortful than the challenging behavior it will replace. Second, practitioners should consider the likelihood that the response will be recognized in the child's daily life. Third, practitioners should consider the likelihood that the individual will acquire the selected communication quickly (Tiger et al., 2008).

Interventionists must select a context for implementing FCT, including an implementer and location (Tiger et al., 2008). FCT has been studied in a range of contexts, from researchers implementing FCT in highly controlled clinical settings to

parents implementing FCT in community settings (Tiger et al., 2008). While this review focuses on parent-implemented FCT, researchers or practitioners hoping to train parents as implementers have several considerations related to both interventionist and location. Researchers or clinicians can implement FCT initially and then introduce parents (e.g., Casey, Perrin, Merical, Lecomte, Milligan, & Walsh-Czekalski, 2008) or parents can implement FCT from the beginning (e.g., Simacek, Dimian, & McComas, 2017). Additionally, parents can implement FCT in a controlled context or in the natural environment. Tiger and colleagues (2008) suggested implementing FCT in a controlled context initially when addressing dangerous behaviors, but conducting FCT in natural environments with non-dangerous behaviors and after achieving behavior reduction with dangerous behaviors.

Several strategies exist for teaching the communication response, including how to utilize motivating operations and how to prompt responding (Tiger et al., 2008). Implementers must consider whether to contrive motivating operations in order to make the reinforcer more valuable or to capitalize on naturally occurring motivating operations. Implementing FCT within the context of daily routines is one strategy for using natural motivating operations (Simacek et al., 2017). Utilizing naturally occurring establishing operations promotes generalization; however, contriving motivating operations may be necessary to generate sufficient teaching opportunities (Tiger et al., 2008). In addition to motivating operations, implementers must also consider how to prompt the communication response (e.g., vocal prompting, physical prompting). Tiger and colleagues (2008) suggested different prompt strategies may be appropriate in different

situations, but regardless of prompt strategy, implementers should plan to fade prompts over time so that children communicate independently.

Implementing FCT typically involves putting challenging behavior on extinction, but some studies arranged other consequences for problem behavior (Tiger et al., 2008). For example, in a natural environment in which consequences are difficult to control (e.g., peer attention in a classroom), challenging behavior may be reinforced during FCT. Some studies have included punishment for problem behavior if extinction failed to produce a sufficient decrease (e.g., Fisher, Piazza, Cataldo, Harrell, Jefferson, & Conner, 1993). Tiger and colleagues (2008) recommended putting problem behavior on extinction, and when extinction is not possible, they recommended ensuring that the communication response receives more reinforcement than challenging behavior.

Finally, implementers must consider thinning reinforcement for the communication response (Tiger et al., 2008). In all studies reviewed by Tiger and colleagues (2008), the communication response was initially reinforced on a continuous reinforcement schedule; that is, the individual received reinforcement following every occurrence of the communication response. However, in the natural environment, it may not be plausible to reinforce communication every time it occurs. Some studies thinned reinforcement for the communication response after initial acquisition (Hagiopian, Boelter, & Jarmolowicz, 2011). Tiger and colleagues (2008) recommended using continuous reinforcement to teach the communication response initially, and then thinning reinforcement.

Tiger and colleagues' (2008) review generated a thorough practice guide that defined the components of FCT and provided recommendations related to each component. However, this review utilized a narrative format and did not examine the extent to which the recommendations moderated the effectiveness of FCT. The present meta-analysis will build on the work of Tiger and colleagues (2008) by coding the characteristics described in their review and examining the extent to which those characteristics moderate effectiveness.

Reviews on FCT

Tiger and colleagues (2008) conducted a review of studies examining FCT in order to develop guidelines for practitioners. In addition to their review described above, a number of other review articles have examined aspects of FCT for different populations (e.g., children with ASD, Mancil, 2006; young children, Durand & Moskowitz, 2015). Mancil (2006) reviewed eight articles with 22 participants in which FCT addressed challenging behavior and communication deficits for children with ASD. They examined participant and intervention characteristics in the included studies. They found that most participants were male, had limited language ability, and engaged in multiple forms of challenging behavior. Additionally, they found that most studies involved multiple assessment types, and communication responses varied across studies. Overall, FCT reduced challenging behavior for children with ASD across ages, behaviors, settings, and behavioral functions. That review focused on FCT for children with ASD, but FCT has been shown to be effective for different populations (e.g., ID, Kurtz, Boelter, Jarmolowicz, Chin, & Hagopian, 2011). Additionally, the Mancil (2006) review

examined FCT across implementers rather than parent-implemented FCT, so no conclusions can be drawn about parent-implemented FCT specifically.

Durand and Moskowitz (2015) examined 21 studies that used FCT to reduce challenging behavior for children ages 1-6. The authors discussed selecting reinforcement schedules, implementing FCT in natural contexts, and implementing FCT with toddlers. The researchers found mixed results among studies that thinned reinforcement for the communication response following acquisition, and suggested researchers consider using auditory or visual signals with young children to cue them that reinforcement will be less frequent. They found that the majority of studies were conducted in clinical settings (Durand & Moskowitz, 2015). Because young children spend more time at home than older children, Durand and Moskowitz (2015) stressed the need for research on FCT implemented at home. In particular, they emphasized the importance of research that occurs in the context of naturally occurring family routines, such as dinner or playtime. Very few studies identified in Durand and Moskowitz's (2013) review examined FCT for children under 3, but the available evidence suggested FCT effectively reduced challenging behavior for toddlers (Durand & Moskowitz, 2015). This review used a narrative format to generate suggestions specific to FCT with young children, and it highlighted emerging areas in FCT research. However, this review only included studies in which every participant met inclusion criteria; thus, the sample may not have reflected all available information on FCT with young children. Additionally, similar to Mancil's (2006) review, Durand and Moskiwitz's (2015) review did not examine parent-implemented FCT specifically.

The reviews described above (Durand & Moskowitz, 2013; Mancil et al., 2006) used a narrative format and, thus, did not calculate effect sizes or examine moderators. To my knowledge, one meta-analysis exists on FCT. Heath, Ganz, Parker, Burke, and Ninci (2015) conducted a meta-analysis on the effectiveness of FCT to reduce challenging behavior, and examined communication, age, and disability as moderators. The researchers located 36 articles published between 1980 and 2011 that evaluated FCT with participants with disabilities, utilized experimental designs, and reported the effects on challenging behavior. For each study, the researchers coded the mode of communicative response (aided or unaided AAC, verbal, or multiple), the participant's age, and the participant's disability. They calculated effect sizes using robust improvement rate difference (IRD), and combined effect size using a fixed-effects model. Overall, the effect size for FCT was 0.86, which is generally considered a large effect (Parker et al., 2009). Verbal responses and aided AAC had higher effect sizes than unaided AAC (0.83, 0.74, and 0.48 respectively). Primary, elementary, and secondary-age participants had higher effect sizes than adults (0.83, 0.76, 0.78, and 0.64 respectively), and participants with ASD had higher effect sizes than participants with ID only (0.79 and 0.64 respectively).

These results indicated FCT effectively reduced challenging behavior overall, with effects that varied from medium to large depending on age, diagnostic status, and intervention type. However, this study has several limitations. This study used a fixed effects model to calculate effect sizes (Heath et al., 2015). Fixed effects models assume that all studies estimate a single population parameter, and are not recommended

(Cumming, 2009, p. 225). Second, this study examined moderators by comparing the effect sizes and error bars for different groups (e.g., diagnoses, communication modalities). This strategy fails to produce direct information about the difference between groups (Cumming, 2012, p. 155), which is the parameter of interest when examining moderators. Additionally, this approach does not enable the examination of continuous moderators or interactions among moderators (Cumming, 2012, Chapter 9). A regression-type approach to moderator analysis produces direct information about moderating variables and enables the examination of continuous moderators and interactions among moderators (Cumming, 2012, Chapter 9). Finally, this study did not examine the effect of FCT on communication, and a number of moderators like pre-intervention communication level, challenging behavior form and function, implementer characteristics, and intervention characteristics were not examined.

The reviews described above examined FCT regardless of implementer rather than examining parent-implemented FCT. Parent implementation may produce different outcomes than clinician or researcher implementation, and variables relevant to parent-implemented FCT, such as parent training and parent fidelity, were not examined in the reviews described above. To my knowledge, one review article examines parent-implemented FCT (Gerow et al., 2018). Gerow and colleagues (2018) completed a systematic review on parent-implemented FCT for children with ASD. The researchers located 26 articles with 78 participants that evaluated parent-implemented FCT to reduce challenging behavior for children with ASD. They coded participant, parent, setting, assessment, and intervention characteristics. Additionally, they coded implementation

fidelity and detailed information about social validity, including: (1) whether the intervention was implemented in a natural context with typical resources, (2) how researchers assessed social validity (e.g., comparison with typically developing pears, clinical significance, and generalization data), and (3) parent satisfaction. Most children in parent-implemented FCT studies were male (81%), most children were between the ages of 3-5 (72%), and most parent implementers were mothers (65%). Most parent-implemented FCT studies took place in participants' homes (85%). Parents were involved in at least some FBA components in 94% of studies, and parents participated in developing the intervention in 38% of studies. Most studies took place in a typical context (95%), but few used typical resources (15%). Among studies that did not meet the typical resources criteria, most involved interaction with a researcher that the family would not be able to access outside of a research context. All studies reported clinically significant goals and outcomes. Twenty-seven percent compared outcomes to typically developing peers, and 58% assessed generalization and/or maintenance. Approximately one-third of studies reported parent opinion. Among studies that reported parent opinion, nearly all rated the intervention positively in terms of feasibility, efficacy, and likelihood of continuing implementation.

Overall, Gerow and colleagues' (2018) review provides descriptive information about parent-implemented FCT studies and detailed information about social validity. However, because this study was a systematic review rather than a meta-analysis, no information was provided on effect sizes of the included studies, and the extent to which coded variables moderated the effectiveness of FCT was not evaluated. Variables such as

demographics and whether the study included typical resources could potentially moderate outcomes, and examining those variables as moderators could yield information about which contexts predict that FCT will be more or less effective. Additionally, this review was restricted to children with ASD. FCT has also been demonstrated to be effective for children with other disabilities, such as intellectual disability (Kurtz et al., 2011). Finally, this review did not provide information on parent-training procedures. Identifying how the researchers trained parents to implement FCT and whether parent-training procedures moderated outcomes could inform parent-training practices and lead to the development of scalable parent-training protocols.

Taken together, previous reviews on FCT demonstrated that FCT effectively produced moderate to large reductions challenging behavior and increases communication for children across ages, disability categories, communication modalities, and behavioral functions. A number of reviews stressed the importance of parent implementation (Durand & Moskowitz, 2015; Heath et al., 2015). However, the previous meta-analysis on FCT did not examine parent implementation (Heath et al., 2015), and the previous review article on parent-implemented FCT did not calculate effect sizes or examine the extent to which participant, implementer, or intervention characteristics moderated outcomes (Gerow et al., 2018). Additionally, the previous review on parent-implemented FCT was limited to children with ASD (Gerow et al., 2018), and no previous review articles on FCT examined parent-training procedures. The present meta-analysis extends the current literature by examining parent-implemented FCT across disability categories; calculating an effect size for parent-implemented FCT studies;

examining the extent to which participant, implementer, and intervention characteristics moderate outcomes; and examining parent-training procedures.

Reviews on Parent Training to Reduce Challenging Behavior

Although I identified only one review article examines parent-implemented FCT (Gerow et al., 2018), a number of review articles have examined parent-implemented interventions that address children's challenging behavior. Reyno and McGrath (2006) analyzed the extent to which demographic variables predict outcomes for 31 studies on parent-training programs designed to reduce children's challenging behavior. The researchers coded the following demographic variables: single parent status, family size, family income, parent education, maternal age, race, barriers to treatment, child behavior severity, adverse parenting practices, maternal psychopathology, maternal depression, marital satisfaction, negative life events, and parenting stress. Additionally, they coded the following participation variables: treatment attendance and source of referral. Among those variables, single parent status, low family income, low parent education, minority group status, barriers to treatment, and negative life events significantly predicted dropout, but all effects were negligible in magnitude. The following variables significantly predicted diminished treatment outcomes: single parent status, large family size, low family income, barriers to treatment, low treatment attendance, being referred by an agency rather than self-referred, high child behavior severity, maternal depression, and negative life events. Results indicated that many variables beyond the child or the training program influence response to parent-training programs, and that research is needed to evaluate matching parent-training programs to demographic variables.

While Reyno and McGrath (2006) focused on demographic and family characteristics, Kaminski, Valle, Filene, and Boyle (2008) conducted a meta-analysis to examine program components related to parent-training effectiveness. They reviewed 77 studies that evaluated the effects of parent-training programs designed to improve behavior of children under age 7. Overall, the authors found a significant, positive effect of parent-training programs for changing both parent and child behavior. The authors coded the presence or absence of the following program delivery variables: curriculum or manual, modeling, homework, rehearsal, role-playing/practice, separate child instruction, and ancillary services. Then, they evaluated whether those variables predicted differences in parent and child behavior. In terms of parent behavior, practicing with their own child significantly predicted more successful programs, and ancillary services significantly predicted less successful programs. In terms of child behavior, practicing with their own child significantly predicted more successful programs.

Reyno and McGrath (2006) and Kaminski and colleagues (2008) analyzed variables predicting parent-training program outcomes across a wide variety of parent-training programs, including programs designed to improve overall parenting rather than those designed to teach parents to implement a specific intervention. Focusing on a specific type of parent-training program may yield different results. Fettig and Barton (2014) reviewed studies in which parents implemented function-based interventions to reduce challenging behavior for children under age 8. They included only studies with the parent as the primary participant, and of the 13 studies that met their inclusion criteria, three involved FCT. For each study, they coded parent, child, and setting

characteristics; dependent variables; parent collaboration; training and support strategies; fidelity; generalization and maintenance; social validity; evidence of experimental control; and the extent to which included studies met quality indicators. Among the 13 included studies, participants ranged from two to eight years old, and the majority (66%) were male. All studies conducted training in the home, and four also conducted training in community settings. All studies reported collaboration with parents, but the extent of collaboration varied across studies, and only five reported parent involvement in developing the behavior intervention. The researchers found that all studies effectively reduced challenging behavior, but only four met quality indicators, limiting interpretation of the results. All studies reported information about pre-intervention parent training; the most common strategies were modeling and practice. Eleven studies reported information on coaching and support following the initial training; the most common strategies were performance-based feedback, modeling, and practice. While Fettig and Barton (2014) do not specify the end date for their search, the most recent article in their review was published in 2011. Parent-implemented FCT is an emerging area of research that has been influenced by changing technology (e.g., telehealth). Fettig and Barton did not report whether articles in their sample involved telehealth, and they did not code variables specifically related to telehealth coaching (e.g., bug-in-ear coaching, feedback based on videos).

Tomlinson, Gore, and McGill (2018) reviewed 20 articles with 113 participants on telehealth to train interventionists to implement ABA procedures. They included articles that involved training a parent, therapist, or teacher to implement any ABA

procedure (e.g., assessments interventions) using telehealth. Their sample included seven articles that involved parent-implemented FCT. In addition to coding demographic variables, training variables, and outcomes, Tomlinson and colleagues (2018) coded characteristics specifically related to telehealth training. They coded type of technology and whether training included live coaching (e.g., real-time communication over a teleconferencing platform), delayed feedback based on videos, and online modules or videos. They found that all studies included some live coaching, but initial training methods varied (e.g., live coaching for the initial training, online modules). Seven studies included sessions in which the implementer conducted sessions independently, without receiving live coaching. Four studies used online modules, and two studies used delayed feedback based on videos. This review provided descriptive information about telehealth training methods, but did not examine relations between training methods and outcomes. Additionally, this review included articles with a variety of implementers and did not focus specifically on parent implementers.

Previous review articles on parent training included a variety of interventions (Kaminski et al., 2008; Reyno & McGrath, 2006; Tomlinson et al., 2018) rather than focusing on programs that train parents to implement a particular intervention. While these reviews yielded information related to demographic variables (Reyno & McGrath, 2006) and training variables (Kaminski et al., 2008; Tomlinson et al., 2018), a review focused on a particular type of intervention may produce different results. FCT involves a specific set of procedures, and examining parent training among parent-implemented FCT studies might generate more specific recommendations than broader reviews.

Fettig and Barton's (2014) review and Tomlinson and colleagues (2018) review included ABA interventions, however their reviews have several limitations in terms of understanding parent-implemented FCT. Neither review calculated effect sizes or examined moderators, so the extent to which the variables that they examined related to outcomes remains unknown. Additionally, both Fettig and Barton (2014) and Tomlinson and colleagues (2018) focused on various interventions rather than FCT specifically. The present review will build upon previous reviews related to parent training for reducing children's challenging behavior by adapting the demographic variables described by Reyno and McGrath (2006), the program variables described by Kaminski and colleagues (2008) and Fettig and Barton (2014), and the telehealth variables described by Tomlinson and colleagues (2018) to evaluate the relevance of those variables to parent-implemented FCT. Given the limitations in the existing reviews on FCT and parent training to reduce children's challenging behavior, this meta-analysis contributes by examining parent-implemented FCT studies to determine the effects on challenging behavior and communication, and by examining participant, intervention, and coaching characteristics in parent-implemented FCT studies.

Chapter 3

Method

The procedures in this meta-analysis were registered with International Prospective Register (Registration # CRD42018100912, Pennington, 2018). PROSPERO publishes protocols for systematic reviews and meta-analyses to increase transparency and minimize reporting bias (Moher, Liberati, Tetzlaff, & Altman, 2009). I registered with PROSPERO prior to data extraction and adhered to the guidelines for meta-analyses proposed by Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA; Moher et al., 2009). A PRISMA checklist is included in Appendix A.

Search Process

The author and another doctoral student in special education with a background in applied behavior analysis conducted a search of the electronic databases of PsycINFO, Medline, and ERIC using the terms (parent*) and ("FCT" OR "functional communication training" OR "mand training" OR "functional analysis AND communication"). The search was conducted in May 2018. Search terms were based on Heath and colleagues' (2015) meta-analysis on FCT with the addition of the "parent" term. We searched for all peer-reviewed papers, dissertations, technical reports, or conference proceedings published in English. The initial search returned 1,743 unique results.

Second, the two coders conducted a title/abstract review of studies returned in the original search. During the title/abstract review, we excluded studies with no reference to FCT or function-based communication intervention in the title or abstract, studies that the title or abstract directly stated were not parent-implemented, and non-empirical articles

(e.g., review articles, qualitative studies). The title/abstract review eliminated 1,599 studies.

Of the 144 articles remaining after the title/abstract review, five non-circulating dissertations could not be located (Hays, 2004; Hill, 1996; Quevedo, 2018; Schmidt, 2011; Sullivan, 2003). All dissertations were compared to peer-reviewed articles by the same author to identify whether they included the same participants based on participant pseudonym; three dissertations that had the same participants as peer-reviewed articles were excluded (Byiers, 2014; Olive, 2017; Simacek, 2017). After eliminating those articles, we conducted a full-text review of the 136 remaining articles to determine whether they met the inclusion criteria described above. Coders included articles if at least one participant in the study met inclusion criteria. The coders excluded 117 articles, and identified 19 articles that met inclusion criteria. Among the excluded articles, 39 did not meet the criterion for experimental design, 59 did not meet the criterion for control phase, 15 did not meet the criterion for intervention, 105 did not meet the criterion for parent implementation, and 21 did not meet the criterion for the dependent variable. Fifty-seven were excluded for more than one category. Additionally, one article was excluded because I could not obtain sufficient information to calculate effect sizes (Wacker, Scheiltz, Berg, Harding, Padilla Dalmau, & Lee, 2017).

Fourth, we reviewed the reference sections of all included articles, as well as the reference sections on previous review articles on FCT (Durand & Moskowitz, 2013; Fettig & Barton, 2014; Falcomata & Wacker, 2013; Hagopian, Boelter, & Jarmolowicz, 2011; Heath et al., 2015; Gerow et al., 2018; Mancil et al., 2006). During the ancestral

search, three articles were identified but excluded because I could not obtain sufficient information to calculate effect sizes (Kurtz, Chin, Huete, Tarbox, O'Connor, & Paclawskyi, 2003; Wacker, Berg, Harding, Barretto, Rankin, & Ganzer, 2005; Wacker, Berg, Harding, Derby, Asmus, & Healy, 1998). These authors were contacted, and one author responded that no data were available. No other authors responded. The ancestral search identified 9 articles for inclusion. Fifth, to locate unpublished data, I emailed the corresponding author of each included article to request any unpublished data sets that meet the inclusion criteria (Faith, Allison, & Gorman, 1996). However, no unpublished data sets were received. Figure 1 shows a PRISMA flow diagram (Moher et al., 2009) for study inclusion.

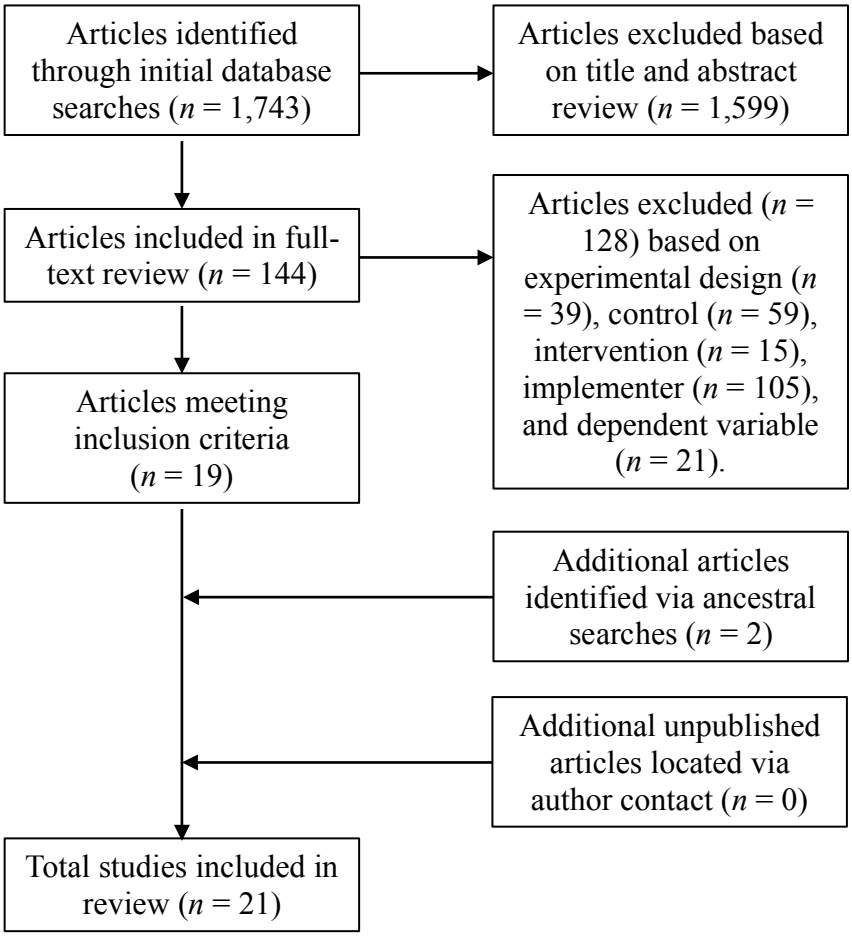

Figure 1. PRISMA flow diagram of search and screening process.

Inclusion Criteria

To examine the effects of parent-implemented FCT for children who engage in challenging behavior, I included studies that met the following criteria. First, the study used an experimental group or single-subject design (e.g., ABAB, multiple baseline, or alternating treatments); quasi-experimental studies were excluded. Although I did not restrict to single-subject studies, no group designs met inclusion criteria, and all included articles were single-subject designs. Second, the study included parent-implemented FCT and a non-treatment, therapist-implemented, extinction, or reversal control group (if

the study was a group design) or phase (if the study was a single-subject design). In order to isolate the effect of parent-implemented FCT, studies that compared two different versions of FCT were excluded (e.g., FR1 and VR3). Third, studies implemented FCT, which consists of an FBA (e.g., structured descriptive assessment or functional analysis) followed by an intervention that teaches a communication response to obtain the functional reinforcer identified in the FBA (Carr & Durand, 1985). Fourth, a parent implemented FCT. I defined parent as any family caretaker (e.g., mother, father, grandparent, legal guardian, foster parent); studies with non-family caretakers (e.g., PCAs) were excluded. Studies that included a training phase in which researchers implemented FCT prior to or in addition to parent implementation were included; studies in which members of the research team, teachers, or other external interventionists were the sole implementers were excluded. Fifth, the study included direct observation of challenging behavior as a dependent variable. I defined challenging behavior using Emerson's (1995) definition: behavior that has the potential to place the safety of the individual or others in serious jeopardy, or behavior that has the potential to limit the individual's access to community facilities. This definition included behaviors likely to harm the individual (e.g., pica, self-injury), behaviors likely to harm others (e.g., aggression), behaviors that challenge caregivers (e.g., noncompliance, tantrums), and behaviors likely to be considered objectionable to the public (e.g., screaming; Emerson, 1995). Studies that reported only prosocial behavior (e.g., the communication response) or that involved increasing variety of vocalizations rather than decreasing challenging behavior were excluded.

Because these studies involved challenging behavior, ethical considerations may compel researchers to limit the number of baseline sessions. Thus, although What Works Clearinghouse recommends at least three data points per phase (Kratochwill et al., 2013), I did not initially exclude studies based on the number of data points per phase. However, after reviewing articles, I excluded one study reporting only one data point in each baseline and intervention phase because it did not contain sufficient data to calculate a meaningful effect size (four data points per participant; Arndorfer et al., 1994). All other studies included at least three data points in the intervention phase.

Search Inter-Rater Agreement

To determine inter-rater agreement (IRA) on the initial search, a second coder conducted the complete initial search on all databases. To determine IRA on inclusion, coders overlapped such that two coders independently evaluated a subset of articles during each phase of the search process. For each phase, the two coders independently screened 5% of articles in that phase to establish reliability. Then, I calculated IRA on that subset of articles. The coders met 90% criterion for the subset, and then they resolved discrepancies and independently screened at least 15% of articles for a minimum of 20% overlap. I calculated IRA on inclusion by dividing the number of agreements (i.e., decision to include or not include an article) by the number of agreements plus disagreements and multiplying by 100. Table 1 shows the percent overlap and percent agreement for each stage of the search process.

Table 1

Inter-Rater Agreement for search process

Search Phase	% overlap	% agreement
Initial Search	100%	97%
Title/Abstract Review	19%	98%
Full Text Review	20%	97%
Ancestral Search	22%	92%

In addition to calculating IRA on inclusion, I calculated IRA on each of the inclusion criteria evaluated in the full-text review. During the full-text review, each coder independently completed a checklist of whether each article met the inclusion criteria. Agreement was defined as both coders coding a criterion as present or absent, and I calculated IRA by dividing agreements by the sum of agreements plus disagreements Table 2 shows IRA on each of the inclusion criteria.

Table 2

Inter-Rater Agreement for inclusion criteria

Criterion	% agreement
Experimental Design	79%
Control Phase	79%
FCT	97%
Parent Implementation	100%
Dependent Variable	97%

Coding Process

To address research questions (RQs) two, four, and five, I used an original coding manual to address participant, implementer, intervention, fidelity, and coaching

information. The author and a second coder independently coded all variables related to RQs two, four, and five at the participant level for every participant who met inclusion criteria. To address RQ seven, I used the CEC standards for evidence-based practices in special education to evaluate the quality and methodological rigor of included studies (CEC, 2014), and the Risk of Bias Assessment for Single-Case Design to evaluate the potential that bias affects study results (Reichow et al., 2015). We coded CEC and Risk of Bias standards at the study level. The following sections summarize coding procedures; see Appendix B for the complete coding manual.

Participant information. To identify participant characteristics (RQ2), I recorded the age, sex, grade, disability, race/ethnicity, and receptive/expressive communication skills of each participant. Additionally, I recorded both the form of the challenging behavior (e.g., aggression, self-injury, etc.) and the function of the challenging behavior as identified in the functional assessment. I coded the function as escape, peer attention, adult attention, tangible, automatic/sensory, or other; if multiple functions were identified, all identified functions were coded.

Parent implementer information. To identify characteristics of parents who implemented FCT (RQ2), I recorded who implemented the intervention and the implementer's age, sex, level of education, socio-economic status, and race/ethnicity. I also coded the following family demographic characteristics described by Reyno and McGrath (2008): single-parent status, family size, low family income, maternal age, minority group status, marital satisfaction, and referral source. I coded adapted versions of the adverse parenting and negative life events/stressors codes described by Reyno and

McGrath (2018), because the codes for adverse parenting and negative life events/stressors involved rating scales not typically reported applied behavior analytic studies. For adverse parenting, I coded whether the authors reported the use of punishment to manage challenging behavior. For negative life events/stressors, I coded whether the authors reported the presence of food insecurity, history of abuse, or history of neglect. Additionally, I coded the number of children present during sessions.

Intervention information. To identify characteristics of FCT (RQ2), I coded the presence or absence of each of the FCT components described by Tiger and colleagues (2008). Additionally, I coded whether the intervention occurred in a natural context with typical resources, whether the parents rated the intervention as feasible and effective, and whether the parents reported that they would continue implementing the intervention based on Gerow and colleagues (2018).

Procedural fidelity. To identify the fidelity with which parents implemented FCT (RQ4), I recorded the level of fidelity reported during baseline and intervention, the fidelity measurement system, the percent of sessions during which fidelity was measured, whether fidelity was distributed evenly within conditions, and whether fidelity was measured on sessions that the parent implemented independently. I also coded whether IOA was calculated on fidelity, and if so, the percent of fidelity sessions with IOA and the level of agreement.

Coaching procedures. To identify the coaching procedures used to train parents to implement FCT (RQ5), I recorded the coach's education, license, training, and years of experience. I coded whether the coach used the family coaching practices described by

Fettig and Barton (2014) or the program delivery variables described by Kaminski and colleagues (2008). Additionally, I coded whether the coach prompted parents during sessions and whether the coach trained parents in didactic/classroom sessions. I coded whether the coach provided training in person and/or over telehealth, and for studies that involved telehealth, I coded the telehealth variables described by Tomlinson and colleagues (2018). Additionally, I coded whether parents met a performance criterion before coaching independently. To examine coaching intensity, I coded the number and length of coaching sessions before, during, and after implementation. I coded whether the researchers specifically described the coach's behavior and measured the coach's fidelity, and if so, I coded the coach's level of fidelity.

Study Quality. To describe the quality of included studies (RQ7) I coded whether the included studies met the CEC (2014) and Risk of Bias Assessment for Single-Case Design (Reichow, Barton, & Maggin, 2015) quality indicators. These quality indicators have been used in a number of recent reviews (e.g., Cowan, Abel, & Candel, 2017; Losinski, Wiseman, White, & Balluch, 2016; Sweigart, Collins, Evanovich, & Cook, 2016). CEC quality indicators were coded as present, partial, or absent (CEC, 2014). Risk of Bias quality indicators were coded as present, absent, or unclear (Reichow et al., 2015). For a complete list of these quality indicators, see Appendix B.

Coding Inter-Rater Agreement

Two coders independently evaluated 29% of included articles with 28% of included participants. For all codes, agreement was scored if both coders selected the

same code for the same participant. I calculated IRA by dividing agreements by agreements plus disagreements and multiplying by 100. First, coders independently evaluated 5% of articles. Then, I calculated IRA on that subset of articles. The coders met 90% criterion for the subset, then they resolved discrepancies then independently screened an additional 24% of articles for a total of 29% overlap. For participant characteristics, IRA exceeded 90% for all codes except for implementer role. IRA for implementer role was 80%. For implementer characteristics, IRA exceeded 90% for all codes. For intervention characteristics, IRA exceeded 90% for all codes except prompt type. IRA for prompt type was 80%. For fidelity, IRA exceeded 90% for all codes. For parent training characteristics, IRA exceeded 90% for all codes except for guided self-reflection. IRA for guided self-reflection was 73%. For CEC quality indicators, IRA was 83% for describing the role of the intervention agent, describing procedures and materials, and fidelity. IRA was 100% for all other codes. For risk of bias, IRA was 100% for sequence generation; blinding of participants, personnel, and outcome assessors; and other sources of bias. IRA was 83% for all remaining codes.

Effect Sizes

To evaluate the effect of parent-implemented FCT on challenging behavior and communication, I calculated two effect sizes for each included participant: one for challenging behavior, and one for communication. All studies located in this meta-analysis were single-subject experimental designs. No consensus exists on the most appropriate effect size for single-subject research, in which data typically fail to meet assumptions of normality and independence (Parker, Vannest, & Davis, 2011). Effect

sizes for single-case data include nonoverlap metrics such as percent nonoverlapping data (PND, nonoverlap of all pairs (NAP), percent of all nonoverlapping data (PAND), Tau-U, and robust improvement rate difference (IRD), and parametric methods like SMD (Parker et al., 2011).

Due to the lack of consensus on an appropriate effect size for single-case research, I calculated multiple effect sizes (e.g., Ledbetter-Cho, O'Reilly, Lang, Watkins, & Lim, 2017; Maggin, Pustejovsky, & Johnson, 2017; Maggin, Zurheide, Pickett, & Bailie, 2015). I calculated PND, PAND, NAP, Robust IRD, PEM, LRR, SMD, and Tau-U. Table 3 describes these effect sizes.

Table 3

Description of effect sizes

Effect Size	Description
Nonparametric	
Improvement Rate Difference (IRD)	The difference in the number of improved data points, or data points remaining after eliminating overlap, in Phase B and Phase A (Parker & Hagan-Burke, 2007).
Nonoverlap of All Pairs (NAP)	Percentage of nonoverlapping pairs of data points out of all possible pairs (Parker & Vannest, 2009).
Percentage of All Nonoverlapping Data (PAND)	The percentage of data points that must be transferred between phases to eliminate overlap, subtracted from 100 (Parker, Hagan-Burke, & Vannest, 2007).
Percent Nonoverlapping	The percentage of data points in Phase B below the lowest

Data (PND)	data point in Phase A for behavior reduction, or above the highest data point in Phase A for behavior increase (Scruggs, Mastropieri, & Casto, 1987).
Tau-U	Nonoverlap between Phases A and B, controlling for baseline trend (Parker, Vannest, Davis, & Sauber, 2011).
Percent Exceeding the Median (PEM)	The percentage of data points in Phase B below the median of Phase A for behavior reduction, or above the median of Phase A for behavior increase (Parker & Hagan-Burke, 2007).
Parametric	
Log Response Ratio (LRR)	Percentage of change in the level of the outcome from Phase A to Phase B (Pustejovsky, 2015).
Standardized Mean Difference (SMD)	The mean difference between phases divided by the standard deviation of Phase A (Hedges,Pustejovsky, & Shadish, 2012).

Calculating multiple effect sizes introduces potential bias by creating the possibility of selecting the effect size that shows the strongest effect to answer the RQ. To counter this risk of bias, I made the a priori decision to answer substantive questions using robust IRD and to provide additional effect sizes as a supplement. I selected robust IRD as my primary effect size calculation for several reasons. First, robust IRD is a non-parametric approach appropriate for single-case data (Parker et al., 2011). Second, robust IRD can be used to produce confidence intervals (Parker et al., 2009). Third, robust IRD is analogous to risk difference in medical research, which makes it recognizable to a

broad audience (Parker et al., 2011). Fourth, criteria exist for interpreting robust IRD (Parker et al., 2009). Parker and colleagues (2009) proposed interpreting robust IRD scores below .50 as small to very small, between .50 and .70 as moderate, and above .70 as large to very large. Fifth, robust IRD was used in a recent meta-analysis on FCT implemented by therapists, teachers, and parents across mode of communication, age, and disability (Heath et al., 2015). Robust IRD can be confounded by positive baseline trend (Parker et al., 2009). Fewer than 5% of studies on FCT in a recent meta-analysis had a positive baseline trend (Heath et al., 2015). We visually analyzed each phase contrast for a positive baseline trend, and found that 5.5% of phase contrasts had a positive baseline trend for challenging behavior, and no phase contrasts had a positive baseline trend for communication.

Single case studies often include multiple phase contrasts for each participant. For example, an ABAB design contains three phase contrasts. Previous meta-analyses have included only the effect size for the first phase contrast (e.g., Heath et al., 2015) or the effect size for every phase contrast. Including only the first effect size is problematic because FCT involves learning, and first phase contrasts may differ systematically from later phase contrasts. Including every phase contrast is problematic if the model fails to account for a lack of independence across effect size observations and for different numbers of effect sizes for different participants. To address these issues, I calculated IRD for every phase contrast, and then I combined effect sizes using a multilevel random effects model that accounted for variability at the participant level. I selected a random effects model because these models assume different studies estimate somewhat different

population effect sizes, which is typically more realistic than assuming a fixed effect size (Cumming, 2012, p. 211). A random effects model is more conservative than a fixed effects model because it allows for greater uncertainty and produces larger confidence intervals (Cumming, 2012, p. 211). Thus, compared to a fixed effects model, a random effects model compromises power, but researchers can be more confident in significant findings from moderator analyses.

Prior to calculating IRD, I extracted the data from each study using Graph Click (Arizona Software Inc., 2010), which has been shown to reliably extract single-case data (Boyle, Samaha, Rodewald, & Hoffmann, 2013). I calculated effect sizes using the SingleCaseES package in r (Pustejovsky & Swan, 2018). I estimated the weighted average effect sizes and generated forest plots using the metafor package in r, which uses maximum likelihood estimation (Viechtbauer, 2010).

Moderator Analysis

To address RQ2, I examined the effects of participant, intervention, and coaching characteristics as potential moderators. Moderator analyses for implementer characteristics were planned but were not conducted due to under-reporting of implementer demographic information. Many single-case meta-analyses use nonparametric tests for moderator analyses (e.g., Mann-Whitney U tests; Walker & Snell, 2013). Although the raw data in a single-case meta-analysis typically fail to conform to the assumptions required for parametric tests, the effect sizes themselves, which are the outcomes in a moderator analysis, can be expected to meet assumptions like independence, normality, and homogeneity of variance required for regression-based

approaches. Regression-based approaches permit examining multiple moderators at a time and examining interactions among moderators (Cumming, 2012, p. 241). Thus, I examined moderators using meta-regression. I used the general linear model with each effect size weighted by its variance.

I examined the following participant characteristics: age, diagnosis, pre-intervention communication level, form of challenging behavior, and function of challenging behavior. I examined the following intervention characteristics: whether a functional analysis was conducted and whether FCT was implemented in the context of natural routines. Additionally, I examined the following coaching characteristics: coaching over telehealth vs. in person, and whether a coach was always present during sessions for which data were reported.

Publication Bias

Studies with significant results may be more likely to be published than studies with null results, and publication bias threatens the validity of a meta-analysis (Sutton, 2009). I included non-peer reviewed articles in my search, and I contacted authors of included articles to locate unpublished data. However, no non-peer reviewed articles were located, and no unpublished manuscripts were obtained. Thus, a moderator analysis planned for study type (e.g., peer-reviewed, unpublished, dissertation) was not conducted. I reported the number of studies with null effects and borderline effects to show how many studies with null effects were published in this area.

Chapter 4

Results

Fifty-three participants from 21 studies met inclusion criteria for challenging behavior, and 29 participants from 14 studies met inclusion criteria for communication. Below, I report the effect of FCT on challenging behavior and communication. Then, I report descriptive characteristics of included studies and the results of moderator analyses. Finally, I report the extent to which included studies met CEC and Risk of Bias quality indicators (Council for Exceptional Children, 2014, Reichow et al., 2015). Table 4 shows characteristics of the included studies.

Table 4

Characteristics of Included Studies

Study	N	Design	Control	Outcome
Anderson et al. (2016)	1	ABAB	Reversal	Both*
Benson et al. (2017)	3	ABAB	Reversal	Both
Berg et al. (2007)	4	ABAB	Extinction	Both
Derby et al. (1997)	4	MBL**	Antecedent	Both
Dunlap & Fox (1996)	6	MBL	Not Reported	CB***
Dunlap et al. (2006)	2	MBL	Business-as-usual	CB
Harding et al. (2009a)	1	ABAB	Extinction	Both
Harding et al. (2009b)	3	MBL	Extinction	Both
Johnson et al. (2004)	1	ABAB	Reversal	Both
Koegel et al. (1998)	3	MBL	Business-as-usual	CB

Lucyshyn et al. (2007)	1	MBL	Business-as-usual	CB
Mancil et al. (2009)	3	MBL	Antecedent	Both
Mancil et al. (2006	1	MBL	Reversal	CB
Olive et al. (2011)	1	MBL	Business-as-usual	Both
Padilla Dalmau et al.	2	Other****	Extinction	Both
(2002)	3	ABAB	Extinction	CB
Scheiltz et al. (2011)	3	MBL	Antecedent	Both
Schindler et al. (2005)	2	Other****	Reversal	Both
Simacek et al. (2017)	2	ABAB	Extinction	Both
Wacker et al. (2011)	3	ABAB	Extinction	Both
Wacker et al. (2013a)	6	ABAB	Extinction	CB
Wacker et al. (2013b)				

*Both communication and challenging behavior

**Multiple Baseline

***Challenging Behavior

****ATD embedded in an ABAB (Padilla Dalmau et al., 2002) and ABAB embedded in a MBL (Simacek et al., 2017)

Effect of Parent-Implemented FCT on Challenging Behavior

The average, weighted overall IRD for the effect of parent-implemented FCT on challenging behavior was 0.72 (SE = 0.03, CI=0.67-0.77). The mean IRD was 0.62 (SD = 0.22). This effect size is at the low end of the large effect range (> 0.7; Parker et al., 2009). Figure 2 shows a forest plot with IRDs and CIs for challenging behavior for each

participant. IRDs ranged from 0.29 (Juan, Schieltz et al., 2011) to 0.95 (David, Mancil et al., 2006). Sixty-one percent of participants had large to very large effects (> 0.70), 28% of participants had moderate effects (0.50 – 0.70), and 11% had small effects (< 0.50). Appendix C shows NAP, PAND, PND, LRR, PEM, and Tau-U effect sizes for challenging behavior.

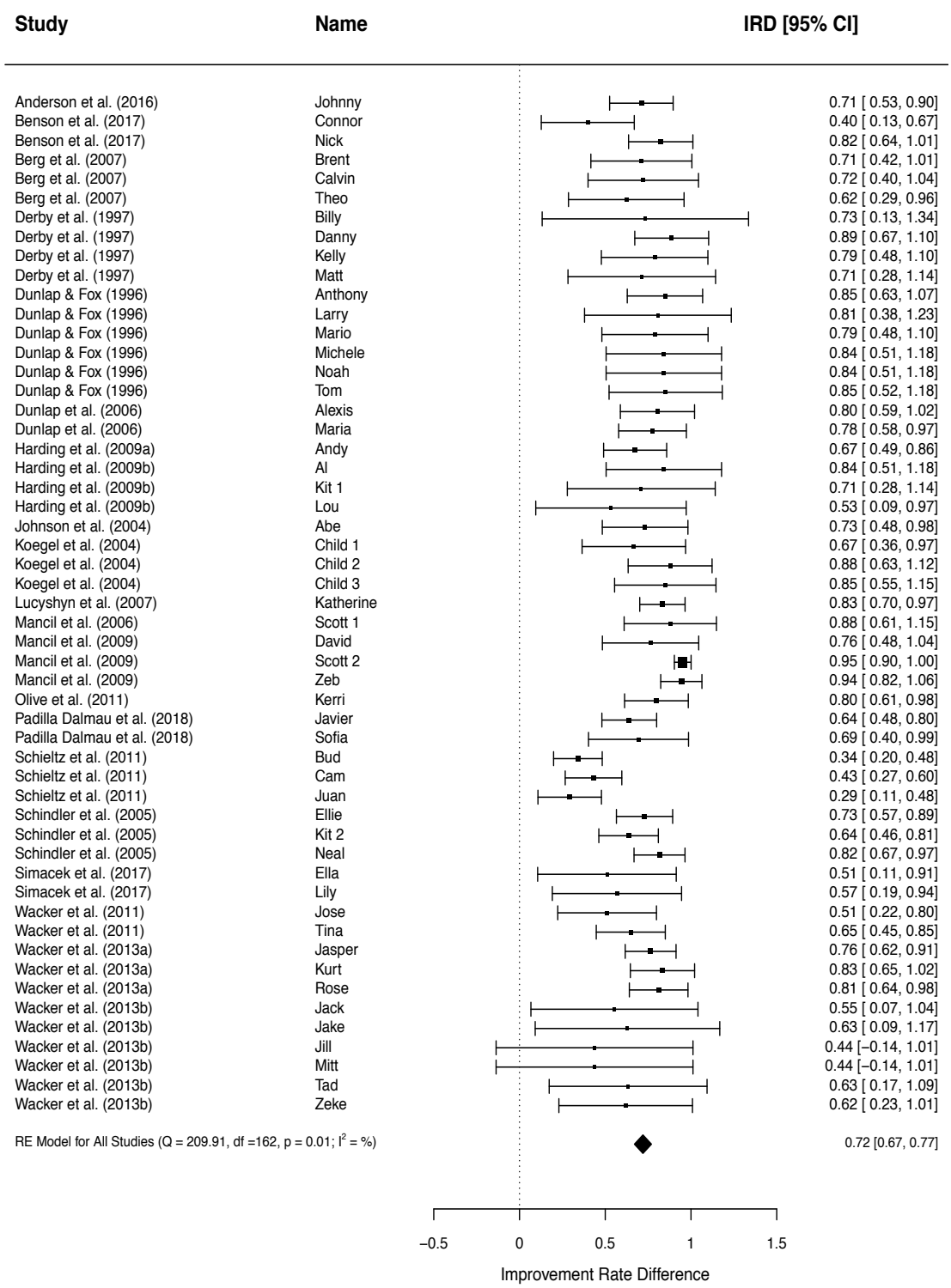

Figure 2. IRD for each participant for effect of FCT on challenging behavior.

Effect of Parent-Implemented FCT on Communication

The average, weighted overall IRD for the effect of parent-implemented FCT on communication was 0.82 (SE = 0.02, CI=0.78-0.85), which is considered a large effect (> 0.70; Parker et al., 2009). The mean IRD for communication was 0.79 (SD = 0.18). Figure 3 shows a forest plot with the IRDs and CIs for communication for each participant. IRDs ranged from 0.39 (Al, Harding et al., 2009b) to 0.93 (Sofia, Padilla Dalmau et al., 2011). Seventy-two percent of participants had large to very large effects (> 0.70), 17% of participants had moderate effects (0.50 – 0.70), and 10% of participants had small effects (< 0.50).

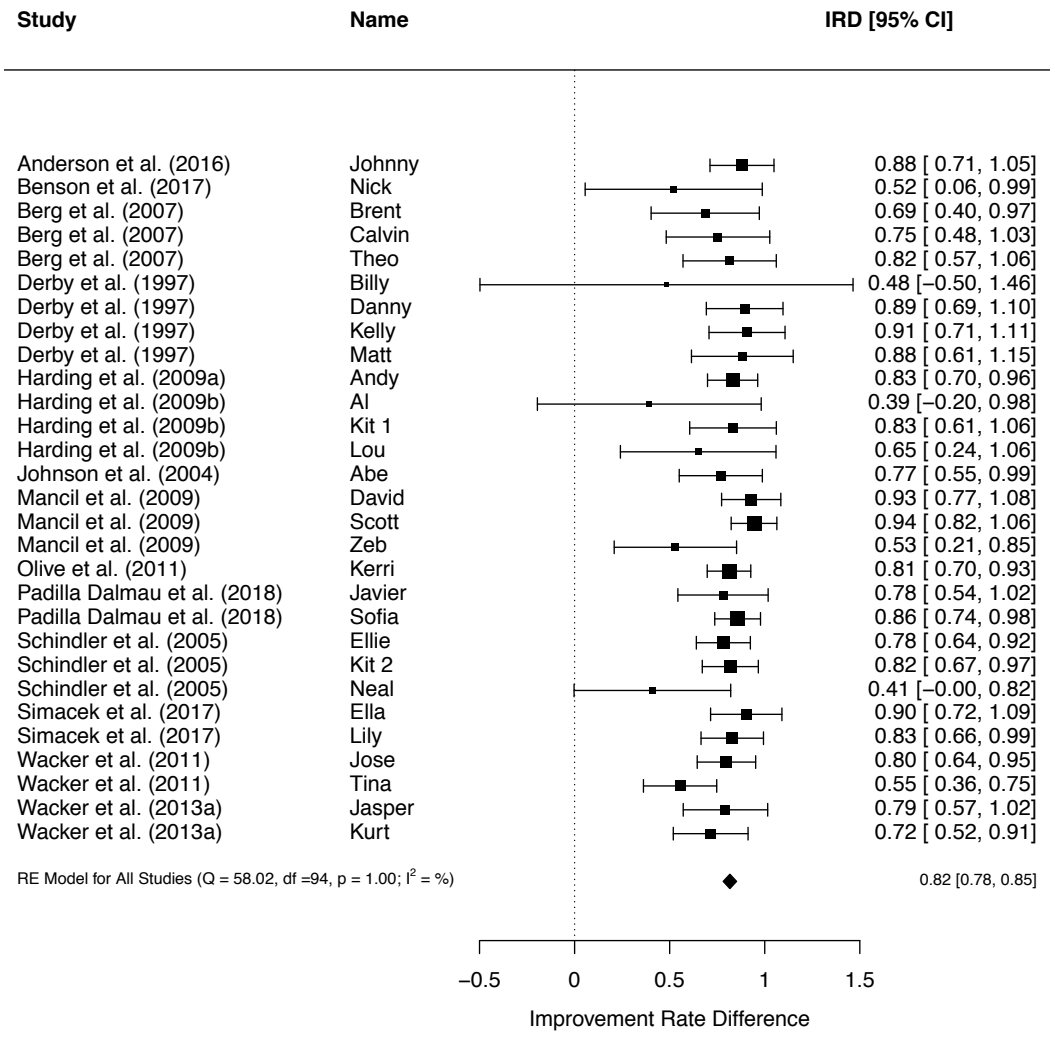

Figure 3. IRD for each participant for effect of FCT on communication.

Participant Characteristics

Participants' ages ranged from 1 year 8 months to 11 years. The majority of participants were male. Most articles did not report race or Hispanic origin. The most commonly reported disabilities were ASD and ID, and eight participants had multiple reported disabilities. Most participants engaged in multiple forms of challenging behavior. The most common behavioral function was determined to be escape, and

nearly half of participants' challenging behavior served multiple functions. Table 5 shows participant characteristics.

Table 5

*Characteristics of Participants in Included Studies (*n=*53)*

Characteristic	***n* (%)**
Age	
< 2	2 (4%)
2	14 (26%)
3	10 (19%)
4	12 (23%)
5-7	13 (25%)
8-10	1 (2%)
11+	1 (2%)
Sex	
Male	40 (75%)
Female	13 (25%)
Race	
Caucasian	5 (9%)
Not Reported	48 (91%)
Hispanic Origin	
Yes	1 (2%)
No	1 (2%)

Not Reported	51 (96%)
Disability	
Autism Spectrum Disorder	33 (62%)
Intellectual Disability	11 (21%)
Developmental Delay	8 (15%)
Speech/Language Delay	3 (6%)
Fragile X Syndrome	1 (2%)
CP	2 (4%)
Charge Syndrome	1 (2%)
Epilepsy	1 (2%)
Spinal Muscular Atrophy	1 (2%)
Visual Impairment	2 (4%)
Multiple Disabilities	8 (15%)
No Disability	1 (2%)
Pre-Intervention Expressive Communication Mode	
None	12 (23%)
Potentially Communicative Acts	4 (8%)
AAC	4 (8%)
Verbal	31 (58%)
Not Reported	6 (11%)
Challenging Behavior Form	

Aggression	31 (58%)
SIB	16 (30%)
Property Destruction	20 (38%)
Tantrums/Vocal Challenging Behavior	20 (38%)
Noncompliance	5 (9%)
Elopement	2 (4%)
Multiple	40 (75%)
Not Reported	6 (11%)
Challenging Behavior Function	
Escape	35 (66%)
Attention	19 (36%)
Tangible	25 (47%)
Automatic	0 (0%)
Multiple	23 (43%)

Moderator analyses for effect on challenging behavior. Based on a previous meta-analysis on FCT that indicated age and diagnosis related to intervention effectiveness, I tested whether age and ASD or ID diagnosis moderated the effect of the FCT on challenging behavior. Additionally, I examined the following moderators that have not previously been examined: presence or absence of pre-intervention communication, engaging in dangerous or non-dangerous challenging behavior, and

behavior function. Age was examined as a continuous moderator; ASD, ID, pre-intervention communication, and dangerous behavior were examined as binary moderators; and behavior function was dummy coded as escape, attention, tangible, or multiple with escape as the reference variable. Based on AIC and BIC, all tested moderator models had worse fit than the reduced model with no moderators. Thus, no further moderator results are reported for the effect of participant characteristics on challenging behavior. Table 6 shows fit statistics for the tested models.

Table 6

Moderator Models for Participant Characteristics Effect on Challenging Behavior

Model	Moderator(s)	Log-Likelihood	Deviance	AIC	BIC
Baseline Model	No moderators	37.2	-74.4	-70.4	-64.22
Model 1	Participant age	37.06	-74.11	-68.11	-58.87
Model 2	ASD + ID	37.04	-74.08	-66.08	-53.78
Model 3	Pre-intervention communication	37.28	-74.55	-68.55	-59.31
Model 4	Aggression/SIB	37.61	-75.23	-69.23	-59.99
Model 5	Attention + Tangible + Multiple Function	37.41	-77.95	-67.95	-52.61

Moderator analyses for effect on communication. I tested whether the same moderators examined for challenging behavior moderated the effect of FCT on

communication. Based on AIC and BIC, all tested moderator models had worse fit than the reduced model with no moderators. Thus, no further moderator results are reported for the effect of participant characteristics on communication. Table 7 shows fit statistics for the tested models.

Table 7

Moderator Models for Participant Characteristics Effect on Communication

Model	Moderator(s)	Log-Likelihood	Deviance	AIC	BIC
Baseline Model	No moderators	35.81	-71.62	-67.61	-62.55
Model 1	Participant age	35.2	-70.4	-64.4	-56.83
Model 2	ASD + ID	35.02	-70.06	-62.05	-52.01
Model 3	Pre-intervention communication	35.21	-70.43	-64.43	-56.86
Model 4	Aggression/SIB	36.31	-72.63	-66.35	-59.99
Model 5	Attention + Tangible + Multiple Function	35.39	-70.79	-60.79	-48.29

Implementer and Family Characteristics

The mother implemented the intervention for 37 participants (70%). For the remaining 14 participants (26%), researchers reported a parent implementer, but it was unclear which parent. Both the mother and father implemented the intervention for one participant (2%). In addition to the parent implementer, a clinician implemented

intervention for one participant (2%), an experimenter implemented intervention for one participant (2%), and a teacher implemented intervention for three participants (4%).

No study reported the implementer's sex, age, or Hispanic status. One study reported that one implementer was of Northern European ancestry. Four studies with eight participants reported the implementer's education level. The implementer had a HS diploma for one participant and a Bachelor's degree for seven participants. Previous ABA training was reported for two of the parents with Bachelor's degrees.

Among the family characteristics related to intervention effectiveness proposed by Reyno and McGrath (2006), no studies reported the presence or absence of the following characteristics: younger maternal age, maternal psychopathology, marital satisfaction, adverse parenting, or negative life events/stressors. Five studies with eight participants reported whether the family was single parent. Two participants had single-parent families and six had two-parent families. Two studies with three participants reported socio-economic status. One participant's family was low-income and two participants' families were not low-income. Two studies with three participants reported minority group status. One participant's family was from a racial or ethnic minority background and two were not. One study with one participant reported a history of maternal depression, and no other studies reported information related to maternal depression. Four studies with 11 participants reported that participants were referred by a medical or educational professional, and no other studies reported referral source.

Three studies with four participants reported family size. Three participants had four-person families, and one participant had a six-person family. Five studies with eight

participants reported whether siblings were present during intervention. Seven participants had one sibling present during intervention, and one participant had no siblings present during intervention. Due to the lack of variability among implementers and low reporting of implementer characteristics, I did not examine implementer characteristics as moderators.

Intervention Characteristics

Over half of interventions were based on the results of a FA, and most interventions used both low-tech AAC and a vocal response as the communication response. Few researchers reported considerations related to Tiger and colleagues' (2008) recommendations, such as response effort and likelihood of recognition; however, researchers determined that response effort matched participant characteristics for most studies and that most responses were likely to be recognized in natural settings. The parents served as the initial implementer in most studies. All studies used prompting, and prompt type varied across studies. Implementers placed challenging behavior on extinction in most studies; in the few studies that involved reinforcing challenging behavior, researchers determined extinction would have been dangerous. All studies used an FR1 schedule to reinforce the communication response. Most studies did not report information related to social validity. Control conditions varied among reversal, extinction, antecedent-manipulation, and business-as-usual conditions. Table 8 shows intervention characteristics.

Table 8

*Intervention Characteristics (*n=*53)*

Characteristic	*n* (%)
Assessment	
Questionnaire/Interview	22 (42%)
Descriptive Assessment	17 (32%)
Structured Descriptive Assessment	7 (13%)
FA	35 (66%)
Communication Response	
High-tech AAC (e.g., electronic device)	0
Low-tech AAC (e.g., PECS)	34 (64%)
Unaided AAC (e.g., gesture/sign)	4 (8%)
Vocal	36 (68%)
Response Effort	
Researchers reported considering response effort	3 (6%)
Coders determined response effort fit participant characteristics	48 (91%)
Likelihood of Recognition	
Researchers reported considering likelihood of recognition	0
Coders determined response was likely to be recognized in natural settings	50 (94%)
Speed of Acquisition	
Researchers reported considering speed of acquisition	0

Initial Implementer	
Researcher	0
Parent	48 (91%)
Teacher/Clinician	3 (6%)
Unclear	2 (4%)
Motivating Operations	
Researchers contrived motivating operations	10 (19%)
FCT occurred in the context of naturally occurring motivating operations	27 (51%)
Prompting	
Prompts used	53 (100%)
Prompts faded	6 (11%)
Prompt Type	
Vocal	15 (28%)
Gestural	11 (21%)
Physical	6 (11%)
Least-to-most	10 (19%)
Most-to-least	2 (4%)
Multiple prompt types	3 (6%)
Not reported	6 (11%)
Consequence for Challenging Behavior	
Extinction	36 (68%)

Punishment	3 (6%)
Reinforcement	6 (11%)
Not Reported	8 (15%)
Reinforcement for Communication	
FR1 reinforcement schedule	53 (100%)
Schedule thinning	0
Demand fading	9 (17%)
Delay to reinforcement	9 (17%)
Conducted with Typically Available Resources	
Yes	8 (15%)
No	42 (85%)
Parent Reported Feasibility	
Yes	9 (17%)
No	0
Not Reported	44 (83%)
Parent Reported Effectiveness	
Yes	25 (47%)
No	0
Not reported	28 (53%)
Parent will Continue to Implement	
Yes	6 (11%)
No	0

Not Reported	47 (89%)

Moderator analyses for effect on challenging behavior. Based on intervention recommendations from previous reviews of FCT (e.g., Tiger et al., 2008), I tested whether the following intervention characteristics moderated the effect of the intervention on challenging behavior: whether an FA was used and whether the intervention occurred in the context of natural routines. FA and natural routines were examined as binary moderators, and control condition was dummy-coded with reversal, extinction, or antecedent manipulation with business-as-usual as the reference variable. Six participants whose control condition was not reported were excluded from the analysis with control condition. Table 9 shows fit statistics for the tested models.

Based on AIC, the model with use of an FA moderating outcomes for challenging behavior had better fit than the baseline model with no moderators. However, the magnitude of the difference was small, and BIC showed worse fit than the baseline model. Therefore, FA was excluded from subsequent models. The model with natural routines had better overall fit than the baseline model according to both AIC and BIC.

Table 9

Moderator Models for Intervention Characteristics Effect on Challenging Behavior

Model	Moderator(s)	Log-Likelihood	Deviance	AIC	BIC

Baseline Model	No moderators	37.2	-74.4	-70.4	-64.22
Model 6	FA	38.79	-77.57	-71.57	-62.32
Model 7	Natural Routines	40.85	-81.7	-75.7	-66.46

Implementing the intervention in natural routines predicted an increase in effect size. Conducting the intervention in the context of natural routines was associated with a 0.14 increase in effect size estimate (SE = 0.05, $p = 0.003$).

Moderator analyses for effect on communication. I tested whether the same moderators examined for challenging behavior moderated the effectiveness of FCT on communication. Based on AIC and BIC, the baseline model with no moderators had better fit than either of the moderator models. Thus, no further moderator results are reported for the effect of intervention characteristics on communication. Table 10 shows fit statistics for the tested models.

Table 10

Moderator Models for Intervention Characteristics Effect on Communication

Model	Moderator(s)	Log-Likelihood	Deviance	AIC	BIC
Baseline Model	No moderators	35.81	-71.61	-67.61	-62.55
Model 6	FA	35.34	-70.68	-64.69	-57.12
Model 7	Natural Routines	35.83	-71.68	-65.67	-58.11

Parent Implementation Fidelity

Researchers collected fidelity data on parent implementation for 14 participants (26%) in six studies, and fidelity levels were reported for 10 participants (19%) in five studies. Among the participants for whom fidelity data were collected, 10 used a checklist and four used percent of correctly delivered consequences. Eleven counted both errors of commission and errors of omission, and three counted only errors of omission.

Among the participants for whom fidelity data were reported, fidelity was measured during baseline and treatment for nine participants and fidelity was only measured fidelity during treatment for one participant. Fidelity was measured during all sessions for eight participants and during 20% of sessions for two participants. Fidelity ranged from 85%-100%. Two studies reported a range of fidelity levels rather than a single level for the study. Of the eight participants with a single reported fidelity level, the mean was 94%. For six participants, fidelity data were collected on sessions the parent implemented independently. For four participants, fidelity data were collected on sessions during which the coach was available to prompt, redirect, or provide feedback to the parent. IOA was collected on parent fidelity for three participants and ranged from 88%-94%.

Parent Training

Researchers trained the parent to implement FCT for 41 participants, clinicians trained the parent for 9 participants, and the coach was unclear for three participants. Two studies with four participants reported demographic data on the coach. The coach

for one study with two participants was a graduate student with three years' experience in applied behavior analysis, and the coach for one study with two participants was a behavior analyst with 20 years' experience. Information about parent training procedures was reported for 45 studies (85%). Table 11 shows parent training characteristics for the 45 participants for whom some training information was reported.

Table 11

Parent Training Characteristics (n=45)

Characteristic	***n* (%)**
Family Coaching Practices	
Intervention self-monitoring	4 (8%)
Collaborative progress monitoring	27 (60%)
Modeling	39 (67%)
Role play	11 (24%)
Practice and support	1 (2%)
Guided self-reflection	9 (20%)
Collaborative problem solving	37 (82%)
Performance-based feedback	35 (78%)
Building motivation	33 (73%)
Program Delivery Variables	
Written manual	35 (78%)
Homework	4 (9%)
Separate child instruction	0

Ancillary services	0
Other Training Techniques	
Prompting	2 (4%)
Didactic instruction	3 (7%)

Coaches conducted all training in person for 43 participants (81%), over telehealth for nine participants (17%), and both in person and over telehealth for one participant (2%). Among the 10 participants whose parent was coached over telehealth, nine involved live videoconferencing during sessions. No studies involved online modules or telephone coaching.

One study (Mancil et al., 2009) with three participants reported a performance criterion that parents had to meet prior to implementing sessions independently. Parents were required to demonstrate correct implementation on 10 consecutive trials. Additionally, parents collected data in that study and were required to achieve 90% IOA before collecting data independently. No other study reported a performance criterion.

Researchers reported session frequency for 44 participants. Four participants had sessions daily, 28 participants had one to four sessions per week, and seven participants had one to three sessions per month. In one study with five participants, researchers reported "weekly to monthly" visits. Researchers reported coaching parents prior to implementation for 45 participants (85%), during implementation for 19 participants (36%), and after implementation for 30 participants (57%). A coach was present in person for every session for 28 participants (53%), and a coach was present over

telehealth for every session for 9 participants (17%). Researchers provided a checklist of coach behaviors for 17 participants (32%), but no study provided fidelity data on coach behavior.

Moderator analyses for effect on challenging behavior. I tested whether the following training characteristics moderated the effect of the intervention on challenging behavior: whether training was conducted over telehealth and whether a coach was present for every session. A coach was considered present for every session if the coach was available to prompt the parent during every intervention session regardless of whether the coach was present in person or over telehealth. Both telehealth and coach presence were coded as binary variables. Based on BIC, all tested moderator models had worse fit than the reduced model with no moderators. The model with telehealth as a moderator had a lower AIC than the baseline model with no moderators; however, the BIC was lower for the baseline model and the difference in AIC between the two models was negligible. Thus, no further moderator results are presented for the effect of training characteristics on challenging behavior. Table 12 shows fit statistics for the tested models.

Table 12

Moderator Models for Training Characteristics Effect on Challenging Behavior

Model	Moderator(s)	Log-Likelihood	Deviance	AIC	BIC

Baseline Model	No moderators	37.2	-74.4	-70.4	-64.22
Model 8	Telehealth	38.5	-77.1	-71.1	-61.8
Model 9	Coach Always Present	26.9	-53.8	-47.8	-38.9

Moderator analyses for effect on communication. I tested whether the same moderators examined for challenging behavior moderated the effectiveness of FCT on communication. Based on AIC and BIC, the baseline model with no moderators had better fit than either of the moderator models. Thus, no further moderator results are reported for the effects of training characteristics on communication. Table 13 shows fit statistics for the tested models.

Table 13

Moderator Models for Training Characteristics Effect on Communication

Model	Moderator(s)	Log-Likelihood	Deviance	AIC	BIC
Baseline Model	No moderators	35.81	-70.61	-67.61	-62.54
Model 8	Telehealth	35.12	-70.24	-64.24	-56.68
Model 9	Coach Always Present	33.82	-67.65	-61.65	-54.29

Quality Indicators

Quality indicators were coded at the study level. Table 14 shows the CEC quality indicators (CEC, 2014). Table 15 shows the Risk of Bias Assessment for Single-Case Design (Reichow et al., 2015). Other sources of bias included two studies that used FA

sessions in which the AAC device did not appear to be present for baseline, and one study that had long gaps between sessions. Overall, most studies at least partially met CEC quality indicators. All studies had at least one indicator of a high risk of bias according to the Risk of Bias Assessment for Single-Case Design (Reichow et al., 2015). Across studies, quality indicators related to fidelity were noticeably absent.

Table 14

*Council for Exceptional Children Quality Indicators (*n=*21)*

Quality Indicator	**Present**	**Partial**	**Absent**
The study describes critical features of the context or setting	1 (5%)	18 (86%)	2 (10%)
The study describes participant demographic variables	2 (10%)	19 (90%)	0
The study describes the disability or risk status and method for determining status	7 (33%)	13 (62%)	1 (5%)
The study describes the role of the intervention agent	0	17 (81%)	4 (19%)
The study describes any specific training required to implement the intervention, and indicates whether the interventionist has achieved them.	1 (5%)	15 (71%)	5 (24%)

The study describes detailed intervention procedures and intervention agents' actions.	2 (10%)	19 (90%)	0
The study describes materials.	8 (38%)	12 (57%)	1 (5%)
The study assesses and reports treatment fidelity using direct, reliable measures.	5 (24%)	1 (5%)	15 (71%)
The study assesses and reports dosage or exposure.	3 (14%)	4 (19%)	14 (67%)
The study assesses and reports treatment fidelity regularly throughout implementation.	5 (24%)	1 (5%)	15 (71%)
The researcher controls the independent variable.	21 (100%)	NA	0
The study describes baseline conditions.	11 (52%)	7 (33%)	3 (14%)
Participants have no or limited access to intervention during baseline.	16 (76%)	5 (24%)	0
The design provides at least three demonstrations of experimental effects at three different times.	21 (100%)	NA	0
Baseline phases include at least three data points.	19 (90%)	NA	0

Outcomes are socially important.	21 (100%)	0	0
The study clearly defines measurement of the dependent variables.	18 (86%)	3 (14%)	0
The study reports the effects of the intervention on all measures of the outcome.	20 (95%)	NA	1 (5%)
Frequency and timing of the outcome measures are appropriate.	19 (90%)	0	2 (10%)
The study provides evidence of adequate interobserver agreement.	19 (90%)	1 (5%)	1 (5%)
The study provides a single-subject graph clearly representing outcome data across all phases for each unit of analysis.	20 (95%)	NA	1 (5%)

Table 15

Risk of Bias Assessment for Single-Case Design (n=21)

Quality Indicator	**Low**	**Unclear**	**High**
Sequence generation method	0	0	21 (100%)
Participant Selection	2 (10%)	16 (76%)	3 (14%)
Blinding of participants and personnel	0	0	21 (100%)
Procedural fidelity	2 (10%)	1 (5%)	18 (86%)
Blinding of outcome assessment	1 (5%)	3 (14%)	17 (81%)
Selective outcome reporting	17 (81%)	0	4 (19%)
Dependent variable reliability	17 (81%)	3 (14%)	1 (5%)
Data sampling	18 (86%)	0	3 (14%)
Other sources of bias	18 (86%)	2 (10%)	1 (5%)

Chapter 5

Discussion

The purpose of this study was to examine the effects of parent-implemented FCT on challenging behavior and communication outcomes. Parent-implemented FCT studies were synthesized using meta-analysis to determine an overall effect size for challenging behavior and communication, and participant and intervention characteristics were examined as moderators using meta-regression. Additionally, studies were coded for quality and risk of bias to evaluate the quality of the research base on parent-implemented FCT and inform future research practices. In this chapter, I summarize and discuss the results presented in Chapter 4 according to each research question. Then, I compare findings with this meta-analysis to findings from a previous meta-analysis on FCT and a previous review article on parent-implemented FCT. Finally, I discuss implications, limitations of this meta-analysis, and suggestions for future research.

What is the overall effect of parent-implemented FCT on challenging behavior and communication for children who engage in challenging behavior?

The overall IRD for the effect of parent-implemented FCT on challenging behavior was 0.72 (CI=0.62-0.77). This is considered a large effect (> 0.70; Parker et al., 2009), but the confidence interval does include IRDs indicating a moderate effect (0.50 – 0.70; Parker et al., 2009). Thus, there is a moderate to large difference in the proportion of improved data points for challenging behavior between baseline and intervention across studies. These results indicate that parent-implemented FCT is moderately to highly effective at decreasing challenging behavior.

The overall IRD for the effect of parent-implemented on communication was 0.82 (CI=0.78-0.85). This effect is considered large (> 0.7; Parker et al., 2009). Thus, there is a large difference in the proportion of improved data points for communication between baseline and intervention across studies. These results indicate that parent-implemented FCT is highly effective at increasing communication.

What are the characteristics of participants, implementers, and interventions in parent-implemented FCT studies? To what extent do these characteristics moderate intervention effectiveness?

Participants. Most participants were young children; 72% were 4 years old or younger. This age level is expected for parent-implemented FCT studies, as young children spend more time at home with parents than do older children (Durand & Moskowitz, 2015). The majority of participants were male. Because over half of participants were diagnosed with ASD, this high proportion of male participants may be related to the higher prevalence of ASD among males (Baio et al., 2014). Over 90% of participants were Caucasian and non-Hispanic, indicating under-representation of minority racial and ethnic groups in this body of literature.

All participants but one had a reported disability, which was expected as children with disabilities are more likely to have communication impairments and engage in challenging behavior (McClintock et al., 2003). Most participants (62%) had ASD. The next most common disability was ID (21%). A variety of other disabilities were represented, including physical and visual impairments. Over half (58%) of participants had some verbal communication prior to intervention, and 23% had no communication

prior to intervention. The rest used AAC or idiosyncratic communication. These findings indicate that this literature represents a range of disabilities and pre-intervention modes of communication.

Most participants (75%) engaged in multiple forms of challenging behavior. Over half engaged in aggression, and over 30% engaged in SIB, property destruction, and tantrums. Less common forms of challenging behavior included noncompliance and elopement. Eighty-one percent of participants engaged in at least one form of challenging behavior that posed an immediate physical danger to himself/herself or others. Overall, this findings indicates that this body of literature represents challenging behaviors that are difficult to manage. Escape was the most common challenging behavior function (66%), but attention and tangible functions were also represented (36% and 47% respectively) and 43% of participants' challenging behavior served multiple functions. No participants had automatically maintained challenging behavior, which is expected because FCT addresses communicative functions.

I tested age, ASD or ID diagnosis, the presence or absence of pre-intervention communication, engaging in aggression and/or SIB, and behavioral function as moderators for effects on challenging behavior and communication. All moderator models had worse fit than the baseline model with no moderators. This indicates that parent-implemented FCT had similar effects on challenging behavior and on communication across a variety of ages, diagnoses, pre-intervention communication modes, behavioral forms, and behavioral functions.

Implementers. The implementer's role was reported for 72% of participants. For implementers whose role was reported, 97% were mothers. For 3% of participants, both the mother and the father implemented the intervention. No studies reported the father was the sole implementer.

Few studies reported implementer characteristics. Only 17% of participants had any demographic information about the implementer reported. Due to the low reporting of implementer characteristics, it is difficult to draw conclusions about the sample of implementers in parent-implemented FCT studies. No moderators were examined, so the extent to which implementer characteristics moderate parent-implemented FCT effectiveness remains unknown.

Interventions. Parents implemented FCT initially for nearly all participants, and no studies reported that the researcher was the initial implementer. Over half of interventions were based on the results of multiple assessments, and the most common type of assessment was an FA (66%). This result is expected based on the previous recommendation that implementers use an FA to analyze severe or dangerous problem behavior (Tiger et al., 2008).

The most common type of communication response targeted for intervention was vocal (68%). This finding is expected based on the percentage of participants who had some vocal communication prior to intervention. For 40% of participants, multiple types of communication responses were reinforced. Often, this was because participants who had not previously engaged in vocal behavior and who had an AAC device spontaneously

began emitting vocal responses during intervention, and the vocal responses were also reinforced.

To select a communication response, Tiger and colleagues (2008) recommended interventionists consider the response effort of the communication response and the likelihood that the response would be recognized in natural settings. No researchers reported considering either the response effort or the likelihood of recognition. However, coders determined that the response effort was appropriate and the response was likely to be recognized for nearly all participants. All participants for whom coders determined a mismatch existed between response effort and participant characteristics engaged in vocal behavior prior to intervention but used a microswitch during intervention. All participants who coders determined engaged in a response unlikely to be recognized in natural settings involved an idiosyncratic sign or gesture (not the standard ASL sign).

Researchers contrived motivating operations to increase the value of the reinforcer for 19% of participants, and FCT was conducted in a setting where the motivating operation naturally occurred for 51% of participants. The remaining studies did not report considering or utilizing motivating operations. The interventionist prompted the communication response for all participants, and prompt types varied. This finding is expected based on previous recommendations to prompt the appropriate communication response (Tiger et al., 2008). Prompt fading, which was recommended by Tiger and colleagues (2008), was reported for 11% of participants. This indicates that this sample of literature included various prompt types, but prompt fading has not been widely examined and no conclusions about prompt fading can be drawn.

The most commonly reported consequence for challenging behavior was extinction (68%), which is consistent with recommendations to put challenging behavior on extinction during FCT in most situations (Tiger et al., 2011). All studies reinforced the communication response on an FR1 schedule, and none reported thinning the schedule of reinforcement. Thus, no conclusions can be drawn from this body of literature about schedule thinning. Tiger and colleagues (2008) recommend an FR1 schedule for initial acquisition, but because communication is rarely reinforced on every occurrence in the natural environment, they recommend thinning to a more natural reinforcement schedule after acquisition.

Most studies (85%) were conducted with access to the researcher, who would not typically be available outside of the study. The 15% of studies coded as using typically available resources used a clinician or teacher as the coach. The majority of studies did not report any social validity data. Among the studies that did, all reported that parents found the intervention feasible and effective, and all reported that the parents would continue to implement.

I tested the use of an FA and conducting the intervention in the context of naturally occurring routines as moderators for challenging behavior and communication. For challenging behavior, results did not indicate that using an FA moderated effect sizes. However, for natural routines, results indicated that the moderator model predicted effect sizes fit the data better than the baseline model with no moderators. Implementing the intervention in the context of natural routines was associated with an increased effect size, indicating that implementing the intervention in the context of natural routines

related to greater differentiation between levels of challenging behavior in baseline and intervention. This finding is encouraging because it indicates that conducting intervention in natural routines, which is associated with a higher likelihood of generalization (Steege, Mace, Perry, & Longenecker, 2007), also produces a greater initial effect. For communication, the results did not indicate that any tested intervention characteristics moderated communication outcomes.

To what extent do parents implement FCT with fidelity?

No fidelity data were reported for the majority of participants (81%). For the 19% of participants for whom fidelity data were reported, the parent's implementation fidelity ranged from 85%-100%. Among those participants, fidelity data were collected when the coach was available to prompt or redirect the parent for 40% of participants, and fidelity data were collected only when the parent implemented independently for 60% of participants. Due to low reporting, conclusions cannot be drawn about the extent to which parents implement FCT with fidelity.

How were parents coached on implementation? To what extent do coaching characteristics relate to intervention effectiveness?

Researchers trained parents to implement FCT for most participants (77%). Only four participants had demographic data reported on the coach; thus, no conclusions can be drawn about coach demographic characteristics in this sample. Information about coaching procedures was reported for 85% of participants.

Among the participants for whom parent-coaching procedures were reported, most used multiple coaching practices. The most common practices were progress

monitoring, modeling, collaborative problem solving, performance-based feedback, and building motivation. Most studies provided the parents with a written manual.

Coaches were trained exclusively in person for most participants (81%). For the remaining participants, all or some training was conducted over telehealth. Coaching frequency was difficult to evaluate, as it was commonly reported in ranges (e.g., 1-5 times per week, 15-30 minute sessions). Thus, coaching intensity could not be evaluated. A coach was present in person or over telehealth for every session for 70% of participants. Therefore, in the moderator analysis, coaching intensity was examined as whether or not the coach was always present during sessions.

I tested conducting coaching over telehealth and whether a coach was always present as moderators for challenging behavior and communication. All moderator models had worse fit than the baseline model with no moderators. This indicates that parent-implemented FCT had similar effects on challenging behavior and on communication regardless of whether training was conducted over teleheath or the coach is always present.

To what extent do included studies meet CEC and Risk of Bias quality indicators?

Overall, most studies partially met CEC (2014) quality indicators for describing the context, setting, participants, and intervention procedures. Most studies did not meet quality indicators for fidelity. Most studies met quality indicators for internal validity, outcome measures, and data analysis. For CEC (2014) to classify a practice as evidence-based or potentially evidence-based, it must be supported by methodologically sound studies showing positive effects. For CEC (2014) to classify a practice as having

negative effects, methodologically sound studies must show negative effects. Methodologically sound is defined as meeting all quality indicators (CEC, 2014). No studies included in this review met all quality indicators. Therefore, according to the CEC quality indicators, parent-implemented FCT has insufficient evidence to meet the criteria for a determination.

All studies had high risk of bias (Reichow et al., 2015) in at least two categories. Overall, risk of bias was high for 100% of studies for sequence generation method and blinding of participants and personnel, and for over 80% of studies for procedural fidelity and blinding of outcome assessment. The risk of bias was low for over 80% of studies for selective outcome reporting, dependent variable reliability, data sampling, and other sources of bias. These findings indicate that researchers this sample of studies may be biased by researchers selecting when to implement phase changes, participants and personnel knowing when the intervention is in place, poor fidelity reporting, and data collectors knowing the condition while coding the outcome.

Because no studies met the CEC (2014) methodologically sound criterion and all studies had high risk of bias for at least two indicators (Reichow et al., 2018), these findings should be interpreted with caution. Due to poor fidelity reporting, the integrity of the independent variable in this literature remains largely unknown. No study randomized phase changes, all studies informed the participants and personnel of phase changes, and most studies used coders who were aware of the phase while coding. Thus, researcher bias could influence these results. Future researchers should report detailed

information about procedural fidelity and should consider randomizing phase changes and using coders who are blind to the study phase.

Comparison with Previous Research

A review of the literature located a number of narrative and systematic reviews of FCT (Durand & Moskowitz, 2015; Gerow et al., 2018; Mancil, 2006; Tiger et al., 2008), but only one meta-analysis of FCT (Heath et al., 2015). The present study extends the current literature by using meta-analytic methodology to examine parent-implemented FCT. Below, I discuss how findings from the present study relate to findings from the previous meta-analysis (Heath et al., 2015).

Heath and colleagues (2015) conducted a meta-analysis on the effectiveness of FCT to reduce challenging behavior. Overall, Heath and colleagues (2015) found a large effect size (IRD = 0.89). We also found an effect size that is considered large (IRD = 0.72), but the confidence interval included moderate effects (CI = 0.67 - 0.77). This difference could indicate differences in the effectiveness of FCT broadly and parent-implemented FCT, but it could also be related to methodological differences. Heath and colleagues (2015) used a fixed-effects model to combine effect sizes and the present study used a random-effects model. Alternately, these different findings could be related to differences in the samples for the two articles. Seventy-two percent of participants in the present study were below age 5. This is a higher amount than Heath and colleagues' (2015) meta-analysis, in which 28% of participants were younger than age 5.

Heath and colleagues (2015) found a relation between age and effect size, with lower ages associated with higher effect size. By contrast, the present study did not find

that age moderated effect size. This disparity could exist for several reasons. First, the studies used different methods to examine moderators. Heath and colleagues (2015) examined age as a moderator by comparing error bars for different age categories, which does not directly test the difference between ages and can lead to overestimating between-group differences (Cumming, 2012, p. 155). I tested age as a continuous moderator using meta-regression. Second, as noted above, participant ages were differently distributed between the two studies, with Heath and colleagues' (2015) study including older participants. Thus, age might moderate effect sizes among a more age-varied sample but not among the sample in the present study. Third, Heath and colleagues (2015) examined all studies on FCT, while the present study included only studies with parent implementers. It is plausible that age interacts differently with effect size when parents are implementers.

Similarly, Heath and colleagues (2015) found that FCT was more effective for individuals with ASD than for individuals with ID. This study did not find a relation between diagnosis and effect size. These disparate findings could be due to different methods or to different relations between diagnosis and effect sizes among parent implementers.

Heath and colleagues (2015) did not examine the following characteristics that were examined in the present review: pre-intervention communication, form of challenging behavior, function of challenging behavior, implementer characteristics, or intervention characteristics. Additionally, Heath and colleagues (2015) did not examine the effect of parent-implemented FCT on communication. Thus, the present meta-

analysis extends the meta-analysis by Heath and colleagues (2015) by focusing on parent-implemented FCT, reporting effects on communication, using a random effects model rather than a fixed effects model to combine effect sizes, and using meta-regression to examine moderators. Additionally, compared to Heath and colleagues (2015), the present review describes additional participant characteristics; describes implementer, intervention, and coaching characteristics; and examines additional moderators.

Limitations and Future Directions

Despite including non-peer reviewed sources in the search and contacting authors to request unpublished manuscripts, only peer-reviewed articles were located. This study found that 11% of participants experienced small effects and 4% of had a confidence interval that included negative effects, indicating that some studies with null effects are published. However, the extent to which publication bias affects this body of literature remains unknown, and different results may have been found if the sample included non-peer-reviewed sources. Future researchers should examine non-peer reviewed sources.

Although effect sizes varied across participants, the present study identified only one moderator related to intervention effectiveness, conducting the intervention in the context of natural routines. Future researchers should examine additional potential moderators. Identifying characteristics that moderate effect sizes could inform decisions about when parent-implemented FCT is most appropriate and could refine intervention practices.

Aspects of communication like pre-intervention communication level or how the child used functional communication prior to intervention might influence intervention

results. The present study considered only the pre-intervention mode of communication. Thus, children's pre-intervention communication level and how they used communication prior to intervention remain unknown. Future meta-analysts should code pre-intervention level of receptive and expressive communication and whether or how the child used communication prior to intervention.

Few studies reported implementer characteristics. The sample of implementers in this study cannot be described, and the extent to which implementer characteristics moderate outcomes could not be examined. Future researchers conducting parent-implemented FCT should report thorough information about the implementers. Similarly, fidelity data were underreported and no conclusions can be drawn about the extent to which parents implement FCT with fidelity. Future researchers should report detailed fidelity data so that future meta-analysts can evaluate parent's overall level of fidelity and examine relations between fidelity and outcomes.

A number of intervention characteristics suggested by Tiger and colleagues (2008) as potentially related to FCT effectiveness have not been examined in parent-implemented FCT studies or have been examined infrequently. For example, no studies reported schedule thinning, and only one study reported prompt fading. Future researchers should consider fading prompts and thinning the reinforcement schedule for the communication response.

Applied Implications

The present findings indicate that parent-implemented FCT reduced challenging behavior and increased communication regardless of the child's age, diagnosis, pre-

intervention communication, behavioral form, or behavioral function. Additionally, parent-implemented FCT reduced challenging behavior regardless of whether an FA was used, whether parent training was conducted over telehealth, and whether a coach was always present. These are promising findings, because they indicate that parents can be effective implementers of an evidence-based behavior intervention for their own children. The present study found that parent-implemented FCT was more effective when implemented in natural settings. This is encouraging because implementing FCT in natural settings is recommended and more likely to produce generalizable results (Tiger et al., 2008). Practitioners should be encouraged to train parents to implement FCT and to consider training parents to implement FCT in natural settings, and they should feel comfortable training parents using telehealth.

Conclusion

This study investigated the effect of parent-implemented FCT on decreasing challenging behavior and increasing communication for children who engage in challenging behavior. This study used a random effects meta-analysis model to combine effect sizes and examined moderators using mixed-effects meta-regression. Findings indicate that parent-implemented FCT decreased challenging behavior and increased communication across a variety of participant and intervention characteristics.

Due primarily to underreporting of demographic and fidelity information, no studies met criteria for methodologically sound studies according to CEC (2014) quality indicators. Similarly, all studies had some risk of bias (Reichow et al., 2018). However, as more research is conducted to address these issues, practitioners should feel

encouraged. Overall, these findings support training parents as interventionists, which is a promising strategy due to the shortage of service providers and the lengthy waitlists for behavior intervention services (Betz et al., 2012).

References

Anderson, C., & Long, E. (2002). Use of a structured descriptive assessment methodology to identify variables affecting problem behavior. *Journal of Applied Behavioral Analysis, 35,* 137–154. doi: 10.1901/jaba.2002.35-137

Arizona Software Inc. (2010). GraphClick 3.0.2. (http://www.arizona-software.ch/graphclick/)

Auerbach, C. & Zeitlin, W. (2018). SSDforR: Functions to analyze single system data. R package version 1.4.18. https://CRAN.R-project.org/package=SSDforR

Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis*, *1*, 91–97.

Baio, J.B, Wiggins, L., Christensen, D.L., Maenner, M.J., Daniels, J., Warren, Z., …Dowling, N.F. (2014). Prevalence of autism spectrum disorder among children aged 8 years – Autism and Developmental Disabilities Monitoring Network, 11 sites, United States. Retrieved from https://www.cdc.gov/mmwr/volumes/67/ss/pdfs/ss6706a1-H.pdf

Beitchman, J. H., Wilson, B., Brownlie, E. B., Walters, H., & Lancee, W. (1996). Long-term consistency in speech/language profiles: I. Developmental and academic outcomes. *Journal of the American Academy of Child and Adolescent Psychiatry*, *35*, 804–814. http://doi.org/10.1097/00004583-199606000-00021

Betz, C. L., Baer, M. T., Poulsen, M., Vahanvaty, U., Bare, M., Haddad, Y., & Nwachuken, G. (2004). Secondary analysis of primary and preventive services accessed and perceived service barriers by children with developmental

disabilities and their families. *Comprehensive Child and Adolescent Nursing*, *27*, 83–106. http://doi.org/10.1080/01460860490451813

Boyle, M. A., Samaha, A. L., Rodewald, A. M., & Hoffmann, A. N. (2013). Evaluation of the reliability and validity of GraphClick as a data extraction program. *Computers in Human Behavior, 29,* 1023–1027. http://doi.org/10.1016/j.chb.2012.07.031

Byiers, B. J. L. (2014). Functional communication training in Rett syndrome. Dissertation Abstracts International Section A: Humanities and Social Sciences, 74(11-A(E)).

Carr, E. G., & Durand, V. M. (1985). Reducing behavior problems through functional communication training. *Journal of Applied Behavior Analysis*, *18*, 1307999. http://doi.org/10.1901/jaba.1985.18-111

Casey, S. D., Perrin, C. J., Merical, C. L., Lecomte, J. M., Milligan, J., & Walsh-Czekalski, M. (2008). Increasing bite acceptance and reducing food refusal in a child with Autism: Moving beyond the clinic. *Journal of Behavior Analysis in Health, Sports, Fitness, and Medicine, 1,* 34–44.

Council for Exceptional Children. (2014). Standards for evidence-based practices in special education. Arlington, VA: CEC.

Cooper, J.O, Heron, T.E., & Heward, W.L. (2007). *Applied Behavior Analysis* (2nd ed). Pearson: New Jersey. ISBN: 0-13-242113-1.

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. http://doi.org/10.1007/s10803-017-3076-6

Cumming, G. (2012). Two simple designs. In Cumming, G., *Understanding the New Statistics: Effect sizes, confidence intervals, and meta-analysis* (pp 153-180). New York: Routledge.

Cumming, G. (2012). Meta-analysis 3: Larger-scale analyses. In Cumming, G., *Understanding the New Statistics: Effect sizes, confidence intervals, and meta-analysis* (pp 153-180). New York, Routledge.

Davis, D. H., Fredrick, L. D., Alberto, P. A., & Gama, R. (2014). Functional communication training without extinction using concurrent schedules of differing magnitudes of reinforcement in classrooms. *Journal of Developmental and Physical Disabilities, 26,* 263–270.

De Luca, R. V, & Holborn, S. W. (1990). Effects of fixed-interval and fixed-ratio schedules of token reinforcement with obese and nonobese boys. *The Psychological Record, 40,* 67-82.

De Luca, R. V, & Holborn, S. W. (1992). Effects of a variable-ratio reinforcement schedule with changing criteria on exercise in obese and nonobese boys. *Journal of Applied Behavior Analysis, 25*, 671–679. http://doi.org/10.1901/jaba.1992.25-671

Dimian, A. F., Botteron, K. N., Dager, S. R., Elison, J. T., Estes, A. M., Pruett, J. R., … Wolff, J. J. (2017). Potential risk factors for the development of self-injurious behavior among infants at risk for autism spectrum disorder. *Journal of Autism*

and Developmental Disorders, *47*, 1403–1415. http://doi.org/10.1007/s10803-017-3057-9

Doughty, A. H., & Lattal, K. A. (2010) Response persistence under variable-time schedules following immediate and unsignalled delayed reinforcement, *Quarterly Journal of Experimental Psychology, 56,* 267-277, DOI: 10.1080/02724990244000124

Duerden, E. G., Oatley, H. K., Mak-Fan, K. M., McGrath, P. A., Taylor, M. J., Szatmari, P., & Roberts, S. W. (2012). Risk factors associated with self-injurious behaviors in children and adolescents with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *42*, 2460–2470. http://doi.org/10.1007/s10803-012-1497-9

Emerson, E. (1995). *Challenging behavior: Analysis and intervention in people with learning disabilities.* New York, NY: Cambridge University Press.

Emerson, E., Cummings, R., Barrett, S., Hughes, H., McCool, C. and Toogood, A. (1988), Challenging behaviour and community services: Who are the people who challenge services?. *Mental Handicap, Vol. 16*, 16-19.

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. http://doi.org/10.1016/S0891-4222(00)00061-5

Faith, M. S., Allison, D. B., & Gorman, B. S. (1996). Meta-analysis of single-case research. In R. D. Franklin, D. B. Allison, & B. S. Gorman (Eds.), *Design and*

analysis of single-case research (pp. 245-277). Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.

Falcomata, T. S., & Wacker, D. P. (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. http://doi.org/http://dx.doi.org/10.1007/s10882-012-9311-3

Fettig, A., & Barton, E. E. (2014). Parent implementation of function-based intervention to reduce children's challenging behavior: A literature review. *Topics in Early Childhood Special Education, 34,* 49–61. http://doi.org/http://dx.doi.org/10.1177/0271121413513037

Fisher, W. W., Piazza, C., Cataldo, M., Harrell, R., Jefferson, G., & Conner, R. (1993). Functional communication training with and without extinction and punishment. *Journal of Applied Behavior Analysis, 26,* 23–36.

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. http://doi.org/10.1177/0145445517740872

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.

Hanley, G. P., Iwata, B. A., & Thompson, R. H. (2001). Reinforcement schedule thinning following treatment with functional communication training. *Journal of Applied Behavior Analysis, 34,* 17-38.

Hays, P. W. (2004). An analysis of functional communication training in a vocational/occupational training high school. Dissertation Abstracts International: Section B: The Sciences and Engineering, 64(Industrial & Organizational Psychology [3600]), 6315.

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. http://doi.org/10.1007/s40489-014-0044-3

Hedges, L. V., Pustejovsky, J. E., & Shadish, W. R. (2012). A standardized mean difference effect size for single case designs. *Research Synthesis Methods*, *3*, 224–239. http://doi.org/10.1002/jrsm.1052

Hernandez, E., Hanley, G. P., Ingvarsson, E. I., & Tiger, J. H. (2007). A preliminary evaluation of the emergence of novel mand forms. *Journal of Applied Behavior Analysis, 40,* 137-156.

Hewitt, A., Timmons, J., Nord, D., Hall-Lande, J., Moore, T., Kliest, B., et al. (2012). A report on early intervention services for Minnesota's children with autism spectrum disorders. Minneapolis, MN: University of Minnesota, Institute on Community Integration, Research and Training Center on Community Living.

Hill, C. (1996). Treating self-injurious and self-stimulatory behaviors with sensory integration: A comparison with communication training. Dissertation Abstracts International: Section B: The Sciences and Engineering, 57(1-B), 0698.

Horner, R. H., & Day, H. (1991). The effects of response efficiency on functionally equivalent competing behaviors. *Journal of Applied Behavior Analysis, 24,* 719-732.

Iwata, B. A., Dorsey, M. F., Slifer, K. J., Bauman, K. E., & Richman, G. S. (1982/1994). Toward a functional analysis of self-injury. *Journal of Applied Behavior Analysis, 27,* 197–209.

Johnston, J.M., & Pennypacker, H.S. (1980). *Strategies and tactics for human behavioral research.* Hilldale, NJ: Erlbaum.

Kaiser, A., Cai, X., Hancock, T., & Foster, E. (2002). Teacher-reported behavior problems and language delays in boys and girls enrolled. *Behavior Disorders*, *28*, 23–29. http://doi.org/10.1177/019874290202800103

Kanne, S. M., & Mazurek, M. O. (2011). Aggression in children and adolescents with ASD: Prevalence and risk factors. *Journal of Autism and Developmental Disorders*, *41*, 926–937. http://doi.org/10.1007/s10803-010-1118-4

Kerns, C. D. (1975). Effects of schedule and amount of observed reinforcement on response persistence. *Journal of Personality and Social Psychology, 31*, 983–991. http://doi.org/10.1037/h0076957

Koritsas, S., & Iacono, T. (2012). Challenging behaviour and associated risk factors: an overview (Part I). *Advances in Mental Health and Intellectual Disabilities*, *6*, 199–214. http://doi.org/10.1108/20441281211236643

Kurtz, P. F., Boelter, E. W., Jarmolowicz, D. P., Chin, M. D., & Hagopian, L. P. (2011). An analysis of functional communication training as an empirically supported treatment for problem behavior displayed by individuals with intellectual disabilities. *Research in Developmental Disabilities, 32,* 2935–2942. http://doi.org/http://dx.doi.org/10.1016/j.ridd.2011.05.009

Lalli, J. S., Browder, D. M., Mace, F. C., & Brown, D. K. (1993). Teacher use of descriptive analysis data to implement interventions to decrease students' problem behaviors. *Journal of Applied Behavior Analysis, 2,* 227–238.

Landa, R., & Hanley, G. P. (2016). An evaluation of multiple-schedule variations to reduce high-rate requests in the picture exchange communication system. *Journal of Applied Behavior Analysis, 49*, 388–393. http://doi.org/10.1002/jaba.285

Lattal, K., Reilly, M., & Kohn, J. (1998). Response persistence under ratio And interval reinforcement schedules. *Journal of the Experimental Analysis of Behavior, 70,* 165–183. http://doi.org/10.1901/jeab.1998.70-165

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. http://doi.org/10.1007/s10803-018-3573-2

Lerman, D. C., & Iwata, B. A. (1996). Developing a technology for the use of operant extinction in clinical settings: an examination of basic and applied research. *Journal of Applied Behavior Analysis, 29,* 345–385. http://doi.org/10.1901/jaba.1996.29-345

L & M Policy Research, LLC. (2014). Autism spectrum disorders: State of the states of services and supports for people with ASD (No. HHSM-500-2006-00009I/HHSM-500-T0002). Washington, DC: Centers for Medicare and Medicaid Services. Retrieved from http://www.autismspeaks.org/sites/default/files/docs/gr/asd-state-of-the-states-report.pdf

Losinski, M., Wiseman, N., White, S. A., & Balluch, F. (2016). A meta-analysis of video-modeling based interventions for reduction of challenging behaviors for students with EBD. *Journal of Special Education, 49,* 243–252. http://doi.org/10.1177/0022466915602493

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. http://doi.org/http://dx.doi.org/10.1177/0741932517716900

Maggin, D. M., Zurheide, J., Pickett, K. C., & Baillie, S. J. (2015). A Systematic Evidence Review of the Check-In/Check-Out Program for Reducing Student Challenging Behaviors. *Journal of Positive Behavior Interventions, 17,* 197–208. http://doi.org/10.1177/1098300715573630

Matson, J. L., & Shoemaker, M. (2009). Intellectual disability and its relationship to autism spectrum disorders. *Research in Developmental Disabilities*, *30*, 1107–1114. http://doi.org/10.1016/j.ridd.2009.06.003

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.

McClintock, K., Hall, S., & Oliver, C. (2003). Risk markers associated with challenging behaviours in people with intellectual disabilities: a meta-analytic study. *Journal of Intellectual Disability Research*, *47*, 405–416. http://doi.org/10.1046/j.1365-2788.2003.00517.x

Metzger, B. A., & Lattal, K. A. (1998). Variable-interval schedule control following response acquisition with delayed reinforcement. *Psychological Record, 48,* 685–696. http://doi.org/10.1037/0097-7403.16.1.27

Merlin, T., Weston, A., & Tooher, R. (2009). Extending an evidence hierarchy to include topics other than treatment: revising the Australian levels of evidence. *BMC Medical Research Methodology, 9*, 1-8.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS med*, *6*, e1000097.

Murphy, G. (2009). Challenging behavior: A barrier to inclusion? *Journal of Policy and Practice in Intellectual Disabilities*, *6*, 89–90. http://doi.org/10.1111/j.1741-1130.2009.00216.x

National Autism Center. (2015). Findings and conclusions: National standards project, phase 2. Randolph, MA. Retrieved from http://www.nationalautismcenter.org/national-standards-project/phase-2/

Olive, Melissa Lynn. (2000). Functional Communication Training: An assessment of stimulus and response generalization. Dissertation Abstracts International Section A: Humanities and Social Sciences, 60(11-A), 3905.

O'Neill, R.E., Albin, R.W., Storey, K., Horner, R.H., & Sprague, J.R. (2015). *Functional assessment and program development for challenging behavior: A practical handbook.* Stamford, CT: Cenage Learning.

Parker, R. I., & Hagan-Burke, S. (2007). Median- based overlap analysis for single case data. *Behavior Modification, 31*, 919–936. doi:10. 1177/0145445507303452

Parker, R. I., Hagan-Burke, S., & Vannest, K. (2007). Percentage of all non-overlapping data (PAND): An alternative to PND. *Journal of Special Education, 40,* 194–204. doi:10.1177/ 00224669070400040101

Parker, R. I., & Vannest, K. (2009). An improved effect size for single-case research: Nonoverlap of all pairs. *Behavior Therapy*, *40*, 357–367. http://doi.org/10.1016/j.beth.2008.10.006

Parker, R. I., Vannest, K. J., & Davis, J. L. (2011). Effect size in single-case research: A review of nine nonoverlap techniques. *Behavior Modification*, *35*, 303–322. http://doi.org/10.1177/0145445511399147

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. doi:10.1016/j.beth. 2010.08.006

Pennington, B. (2018). A meta-analysis on parent-implemented functional communication training to reduce challenging behavior and increase communication: CRD42018100912 Available from: http://www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42018100912

Petscher, E. S., Rey, C., & Bailey, J. S. (2009). A review of empirical support for differential reinforcement of alternative behavior. *Research in Developmental Disabilities, 30,* 409–425. http://doi.org/10.1016/j.ridd.2008.08.008

Pustejovsky, J. E. (2015). Measurement-comparable effect sizes for single-case studies of free operant behavior. *Psychological Methods, 20,* 342–359. doi:10.1037/met0000019

Pustejovsky, J. E. & Swan, D. M. (2018). SingleCaseES: A calculator for single-case effect size indices. R package version rpackageVersion("SingleCaseES"). Retrieved from https://github.com/jepusto/SingleCaseES

Quevedo, N. (2018). An evaluation of schedule thinning progressions involving multiple schedules following functional communication training. Dissertation Abstracts International: Section B: The Sciences and Engineering, 78(9-B(E)), No-Specified.

Qureshi, H., & Alborz, A. (1992). Epidemiology of challenging behaviour. *Mental Handicap Research, 5,* 130–145. http://doi.org/10.1111/j.1468-3148.1992.tb00041.x

Reichow, B. (2012). Overview of meta-analyses on early intensive behavioral intervention for young children with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *42*, 512–520. http://doi.org/10.1007/s10803-011-1218-9

Reichow, B., Barton, E. E., & Maggin, D. (2015). Risk of bias assessment for single case designs. Unpublished manuscript, Anita Zucker Center for Excellence in Early Childhood Studies, University of Florida, Gainesville, FL.

Reyno, S. M., & McGrath, P. J. (2006). Predictors of parent training efficacy for child externalizing behavior problem-a meta-analytic review. *Journal of Child Psychology and Psychiatry*, *47*, 99–111. http://doi.org/http://dx.doi.org/10.1111/j.1469-7610.2005.01544.x

Rose, J. (1999). Stress and residential staff who work with people who have an intellectual disability: a factor analytic study. *Journal of Intellectual Disability Research, Vol. 43*, 268-78.

Schmidt, Jonathan Dean. (2011). Communication assessment and intervention for 3 individuals with autism with severe problem behavior. Dissertation Abstracts International: Section B: The Sciences and Engineering, 71(9-B), 5778.

Simacek, J. J. (2017). Communication intervention for children with severe neurodevelopmental disabilities: An application of telehealth as a service delivery

mechanism. Dissertation Abstracts International Section A: Humanities and Social Sciences, 78(2-A(E)).

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. http://doi.org/http://dx.doi.org/10.1007/s10803-016-3006-z

Skinner, B.F. (1953). *Science and human behavior.* New York: Macmillan.

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. http://doi.org/http://dx.doi.org/10.1007/s10864-013-9183-3

Sullivan, M. T. (2003). Communicative functions of echolalia in children with autism: Assessment and treatment. Dissertation Abstracts International: Section B: The Sciences and Engineering, 63 (Health & Mental Health Treatment & Prevention [3300]), 4948.

Sweigart, C. A., Collins, L. W., Evanovich, L. L., Cook, S. C. (2017). An evaluation of the evidence base for performance feedback to improve teacher praise using CEC's quality indicators. *Education and Treatment of Children, 39,* 419–444. http://doi.org/10.1353/etc.2016.0019

Tiger, J. H., Hanley, G. P., & Bruzek, J. (2008). Functional communication training: A review and practical guide. *Behavior Analysis in Practice*, *1*, 16–23.

Tomlinson, S. R. L., Gore, N., & McGill, P. (2018). Training individuals to implement applied behavior analytic procedures via telehealth: A systematic review of the literature. *Journal of Behavioral Education, 27,* 172–222. http://doi.org/10.1007/s10864-018-9292-0

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*, 1-48. URL: http://www.jstatsoft.org/v36/i03/

Vivanti, G., & Dissanayake, C. (2016). Outcome for children receiving the Early Start Denver Model before and after 48 months. *Journal of Autism and Developmental Disorders, 46,* 2441-2449.

Van den Noortgate, W., & Onghena, P. (2008). A multilevel meta-analysis of single-subject experimental design studies. *Evidence-Based Communication Assessment and Intervention, 2,* 142–151. http://doi.org/10.1080/17489530802505362

Wacker, D.P., Lee, J.F., Padilla Dalmau, Y.C., Kopelman, T.G., Lindgren, S.D., Kuhle, J. ... Waldron, D.B. (2013a). Conducting functional communication training via telehealth to reduce the problem behavior of young children with autism. *Journal of Developmental and Physical Disabilities, 25,* 35-48. doi:10.1007/s10882-012-9314-0

Wacker, D. P., Lee, J. F., Padilla Dalmau, Y. C., Kopelman, T. G., Lindgren, S. D., Kuhle, J., …Waldron, D.B. (2013b). Conducting functional analyses of problem behavior via telehealth. *Journal of Applied Behavior Analysis, 46*, 1–16. doi: 10.1002/jaba.29

Wacker, D. P., Berg, W. K., Harding, J. W., Barretto, A., Rankin, B., & Ganzer, J. (2005). Treatment Effectiveness, Stimulus Generalization, and Acceptability to Parents of Functional Communication Training. *Educational Psychology, 25,* 233–256. https://doi.org/10.1080/0144341042000301184

Walker, V. L., & Snell, M. E. (2013). Effects of augmentative and alternative communication on challenging behavior: A meta-analysis. *AAC: Augmentative and Alternative Communication*, *29*, 117–131. http://doi.org/10.3109/07434618.2013.785020

Wise, M. D., Little, A. A., Holliman, J. B., Wise, P. H., & Wang, C. J. (2010). Can state early intervention programs meet the increased demand of children suspected of having autism spectrum disorders? *Journal of Developmental & Behavioral Pediatrics, 31*, 469–476.

Worcester, J., Nesman, T. M., Raffaele Mendez, L. M., & Keller, H. R. (2005). Giving voice to parents of young children with challenging behavior. *Exceptional Children*, *66*, 509–525.

Yingling, M. E., Hock, R. M., & Bell, B. A. (2017). Time-Lag between diagnosis of autism spectrum disorder and onset of publicly-funded early intensive behavioral intervention: Do race-ethnicity and neighborhood matter? *Journal of Autism and Developmental Disorders*, *48*, 1–11. http://doi.org/10.1007/s10803-017-3354-3

Appendix A

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	4
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	9
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	20
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	40
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	43
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	40
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	40
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	40
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	46
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	46
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used	48

		in any data synthesis.	
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	49
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	49
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	54
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	53
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	55
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	60
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	78
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	58
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	58
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	59
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	61
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	79
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	89
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	89

FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

Appendix B

Coding Manual

Code	Description
Participant Information	
Age	
Sex	
Grade	
Race	
Hispanic Origin	
Disability	
Receptive Communication	Receptive communication was coded as none, followed one-step directions, followed two-step directions, or other.
Expressive Communication	Type of expressive communication was recorded as none, AAC use, potentially communicative acts, or verbal communication.
Dangerous Behavior	Dangerous behavior was coded present if the participant engaged in at least one form of challenging behavior that posted physical danger to self or others.
Aggression	Aggression was coded present if the participant engaged in aggressive behavior directed towards others.
Aggression	Aggression was coded present or absent.

Self-Injurious Behavior	SIB was coded present or absent.
Tantrums/Vocal	Tantrums or other vocal challenging behavior (e.g., yelling, vocal noncompliance) were coded present or absent.
Challenging Behavior Function	The function of challenging behavior was coded as escape, adult attention, peer attention, tangible, automatic, or other. If the authors reported multiple functions, all reported functions were coded.
Number of Children Present	The total number of children present during sessions other than the participant
Implementer Information	
Implementer Role	The role of the implementer (e.g., mother, father) was coded as reported in the study.
Age	
Sex	
Education	
Race	
Hispanic Origin	
Second Implementer	If a second implementer was present, the above information was also coded about the second implementer.

Family Demographic Predictor Variables (adapted from Reyno & McGrath, 2006)

Single-Parent Status	Single-parent status was coded as single-parent family, two-parent family, or not reported.
Family Size	Family size was total number of people in the household, including the participant.
Low Family Income	Low family income was coded present if the authors stated that the family was low income or that the family participated in a program indicating low income (e.g., free/reduced lunch). It was coded absent if the authors stated the family had a middle or high income. Otherwise, it was coded not reported.
Maternal Age	
Paternal Age	
Minority Group Status	Minority group status was coded present if the authors stated the parents were Hispanic or of a racial or ethnic minority. It was coded absent if the authors stated parents did not belong to a minority group. Otherwise, it was coded not reported.
Maternal Psychopathology	Maternal psychopathology was coded present if explicitly reported by the authors. It was coded absent if the authors stated there was no maternal psychopathology. Otherwise, it was coded not reported.

Maternal Depression	Maternal depression was coded present if explicitly reported by the authors. It was coded absent if the authors stated there was no maternal depression. Otherwise, it was coded not reported.
Marital Satisfaction	Marital satisfaction was recorded absent if the authors reported low marital satisfaction. It was recoded present if the authors reported adequate or high marital satisfaction. Otherwise, it was coded not reported.
History of Punishment	Adverse parenting was coded present if the authors reported parents used punishment prior to intervention. It was coded absent if the authors reported parents did not use punishment prior to intervention. Otherwise, it was coded not reported.
Stressors	Stressors were reported present if the authors reported the presence of negative life events or stressors (e.g., food insecurity, history of abuse). It was recorded absent if the authors reported no negative life events or stressors were present. Otherwise, it was coded not reported.
Referral Source	Referral source was coded as self-referred, referred by school or a medical professional, or not reported.

Intervention Information (adapted from Tiger et al., 2008)

Assessment	Assessment type was coded as questionnaire/interview, descriptive assessment, structured descriptive assessment, functional analysis, or other. If multiple assessment types were reported, each assessment type was coded.
Communication Response	Communication response was coded as high-tech AAC, low-tech AAC, unaided AAC, vocal/verbal, or other.
Response Effort Consideration	Coders recorded whether the authors reported considering the response effort of the communication response.
Response Effort Match	Coders recorded whether the response effort of the communication response matched participant information.
Likelihood of Recognition Consideration	Coders recorded whether the authors reported considering the likelihood of recognition of the communication response.
Likelihood of Recognition	Coders recorded whether the communication response was likely to be recognized in natural settings.

Speed of Acquisition Consideration	Coders recorded whether the researchers reported that speed of acquisition of the communication response was considered in response selection.
Initial Implementer	Coders recorded whether the first implementer of FCT was a parent or non-parent (e.g., researcher, teacher).
Motivating Operations	Coders recorded whether researchers contrived motivating operations to increase the value of the reinforcer (e.g., restricting access to parent attention prior to sessions for attention-maintained behavior).
Natural Routines	Coders recorded whether FCT occurred in the context of naturally occurring routines.
Prompting	If prompts were used, coders recorded whether they were verbal, gestural, physical, least-to-most, most-to-least, or multiple prompt types.
Prompt Fading	Coders recorded whether implementers faded the use of prompts.
Extinction	Coders recorded whether the implementer placed challenging behavior on extinction.

Reinforcement	Coders recorded whether reinforcement for challenging behavior was either programmed or expected during FCT (e.g., sibling continued to provide attention). This code did not include errors in fidelity that resulted in reinforcement.
Reinforcer Match	Coders recorded partial match if FCT involved some but not all reinforcers identified in the FBA, complete match if FCT involved all reinforcers identified in the FBA, and overmatch if FCT involved additional reinforcers beyond those identified in the FBA.
Reinforcer	Coders recorded the specific reinforcer used during FCT.
Schedule	Coders recorded the schedule of reinforcement used for the communication response during intervention.
Schedule Thinning	Coders recorded whether the schedule for the communication response was thinned during FCT, and if so they recorded the final schedule.
Demand Fading	Coders recorded whether demand fading was utilized, and if so, they coded the final length of the demand.

Social Validity Information (Adapted from Gerow et al., 2018)

Natural Setting	Natural setting was coded as present if the intervention occurred in a setting where the child would typically be able to access in the absence of research.
Typical Resources	Typical resources was coded as present if the family would have access to materials and trainers in the absence of research.
Feasibility	Feasibility was coded as present if the parent rated the intervention as feasible.
Effectiveness	Effectiveness was coded as present if the parent rated the intervention as effective.
Continue to Implement	Continue to implement was coded Yes if the parent reported they would continue to implement the intervention.

Experimental Design Information

Control Condition	Control condition was coded as either non-parent implemented, reversal, extinction, business-as-usual, or other.
Experimental Design	Experimental Design was coded as ABAB, MBL, ATD, Changing Criterion, or other.

Dependent Measure	The dependent measure was coded for challenging behavior, communication, and any other dependent variables.
Prompted Responses	Coders recorded whether prompted responses were counted

Parent Fidelity Information

Measurement System	Coders recorded the type of fidelity measurement system (e.g., checklist, time sampling) for parent fidelity.
Errors of Commission	Coders recorded whether the fidelity measurement system included errors of commission (e.g., including an extra step).
Errors of Omission	Coders recorded whether the fidelity measurement system included errors of omission (e.g., failure to implement a programmed step).
Baseline	Coders recorded whether researchers measured parent fidelity during baseline.
Treatment	Coders recorded whether researchers measured implementer fidelity during treatment.
Percent of Sessions	Coders recorded the percent of sessions during which fidelity was measured.

Distribution of Sessions	Coders recorded whether the researchers distributed fidelity sessions evenly within conditions.
Independent Implementation	Coders recorded whether parents ever implemented sessions independently without the coach present.
IOA on Fidelity	Coders recorded whether IOA was calculated on fidelity and, if so, what level of IOA was reported.
Coach Demographics	
Coach Role	Role was recorded as reported (e.g., clinician, experimenter)
Education	
License	
Years Experience	
Training	
Family Coaching Practices adapted from Fettig & Barton (2014)	
Self-monitoring	Self-monitoring was coded present if the coach taught parents monitored their own implementation.
Collaborative Progress Monitoring	Collaborative progress monitoring was coded present if the coach worked with parents to create a schedule for monitoring the child's progress.

Modeling	Modeling was coded present if the coach demonstrated some or all of the components of implementing FCT.
Role Play	Role play was coded present if the coach practiced implementing FCT with the parent.
Practice and Support	Practice and support was coded present if the coach provided opportunities for the parent to practice FCT prior.
Guided Self-Reflection	Guided Self-Reflection was coded present if the coach guided the parents to self-reflect on their FCT implementation.
Collaborative Problem Solving	Collaborative problem solving was coded present if the coach worked with the parents to solve any problems that arose with FCT implementation.
Performance-Based Feedback	Performance-based feedback was coded present if the coach provided immediate, specific feedback to parents after FCT sessions.

Telehealth Characteristics, adapted from Tomlinson et al. (2018)

Videoconferencing during Session	Videoconferencing during sessions was coded present if the coach used videoconferencing during sessions while the parent implemented.

Other Videoconferencing	Other videoconferencing was coded present if the coach contacted the parent with videoconferencing other than during sessions (e.g., for meetings prior to sessions).
Online Modules	Online modules was coded present if the coach trained the parent using online training modules.
Telephone Coaching	Telephone coaching was coded present if the coach provided training over the telephone, excluding phone calls related to scheduling sessions.

Other Coaching Characteristics

Prompting	Prompting was coded present if the coach prompted parents during sessions as parents implemented FCT.
Didactic Instruction	Didactic Instruction was coded present if the coach trained parents in didactic/classroom sessions.
Unlisted Coaching Technique	Unlisted coaching techniques were recorded as reported by the authors.
Type of Coaching	Coders recorded whether coaching was delivered in person or with telehealth.

Performance Criterion	If present, coders recorded the performance criterion parents had to meet prior to implementation.
Coaching Intensity	Coders recorded the length and frequency of coaching sessions prior to, during, and after implementation.
Coaching Fidelity	If fidelity information was reported on coach behavior, coders recorded the measurement system and the level of fidelity.

Council for Exceptional Children Quality Indicators (CEC, 2014)

Context and Setting	Coders recorded present if the study described relevant features of the intervention context.
Participant Demographics	Coders recorded present if the study described relevant participant demographic information.
Disability/Risk Status	Coders recorded present if the study described both the participant's disability or risk status and how that status was determined.
Interventionist Role	Coders recorded present if the study described the role of the interventionist and relevant interventionist demographic information.
Interventionist Training	Coders recorded present if the study described specific training or qualifications and whether the interventionist achieved them.

Intervention Procedures	Coders recorded present if the study described intervention procedures with sufficient detail for replication.
Intervention Materials	Coders recorded present if the study described intervention materials with sufficient detail for replication.
Implementation Fidelity	Coders recorded present if the study measured and reports fidelity.
Fidelity Dosage	Coders recorded present if the study reported information related to dosage (e.g., session frequency, length).
Fidelity Pacing	Coders recorded present if the study measured fidelity regularly and across sessions.
Independent Variable	Coders recorded present if the researcher controlled the presence or absence of the independent variable.
Control Condition	Coders recorded present if the study described the control condition with sufficient detail for replication.
Control Condition Access	Coders recorded present if participants had no or very limited access to the intervention in the control condition.
Demonstrations of Effect	Coders recorded present if the design provided opportunity to demonstrate effects at three points in time.

Baseline Phase	Coders recorded present if the baseline phase contained at least three data points.
Experimental Design	Coders recorded present if the design controlled for threats to internal validity (e.g., ABAB, multiple baseline, alternating treatments)
Social Significance	Coders recorded present if outcomes were socially important for the participant.
Dependent Variable	Coders recorded present if the study described the dependent variable and how it was measured sufficiently for replication.
Intervention Effects	Coders recorded present if the study reported the effects on all targeted measures, not only those for which positive effects were found.
Frequency and Timing	Coders recorded present if the study included a minimum of three data points per phase.
Reliability	Coders recorded present if the study reported at least 80% of inter-observer agreement.
IOA Timing	Coders recorded present if the study measured IOA regularly throughout the study and across phases.
Social Validity	Coders recorded present if the study measured social validity.
Data Analysis	Coders recorded present if the study included a single-subject graph that clearly represents outcomes.

Risk of Bias Assessment for Single-Case Design (Reichow et al., 2015)

Sequence Generation	Coders recorded low if the sequence of conditions was selected randomly, high if sequence was not selected randomly, and unclear if randomization was used but the method was unclear.
Participant Selection	Coders recorded low if participants were selected based on the level of target behavior prior to intervention, high if no or inadequate participant selection criteria were reported, and unclear if some participant selection criteria were reported but the level of target behavior prior to intervention was difficult to discern.
Blinding of Participants and Personnel	Coders recorded low if participants, implementers, and individuals making decisions about condition changes were unaware of the condition being implemented; high if no blinding occurred; and unclear if blinding was not reported in sufficient detail to determine.
Procedural Fidelity	Coders recorded low if researchers measured fidelity in at least 20% of sessions across conditions and was at least 80%, high if fidelity was reported for fewer than 20% of sessions or was below 80% in all conditions, and unclear if fidelity was not reported

	across all experimental conditions or one condition had fidelity below 80%.
Blinding Outcome Assessors	Coders recorded low if data collectors were unaware of the condition and purpose of the study, high if coders knew the study condition, and unclear if inadequate information was reported to determine.
Selective Outcome Reporting	Coders recorded low if data were reported for all participants, outcomes, and measures; high if one or more participants did not complete the study or if data were not reported for a participant, outcome, or measure; and unclear if insufficient information were reported regarding attrition or missing data.
Dependent Variable Reliability	Coders recorded low if IOA was at least 80% and was collected for at least 20% of sessions across conditions, high if IOA averaged below 80%, and unclear if IOA was not reported in at least 20% of sessions.
Data Sampling	Coders recorded low if there were adequate data in each condition to determine level and trend, high if there were not sufficient data in each condition, and unclear if data had significant variability.

Other Sources of Bias	Coders recorded any other sources of bias observed in the study.

Appendix C
Effect Sizes for Challenging Behavior

Study	Participant	NAP	PND	PEM	LRR	Tau-U	PAND
Anderson et al. (2016)	Johnny	0.92	0.75	0.95	1.52	0.82	0.89
Benson et al. (2017)	Connor	0.20	0.00	0.00	-0.60	-0.27	0.58
	Nick	0.67	0.00	1.00	2.72	0.50	0.84
Berg et al. (2007)	Brent	0.95	0.95	0.95	1.52	0.89	0.96
	Calvin	0.97	0.93	0.98	2.16	0.94	0.96
	Theo	0.73	0.04	1.00	2.54	0.41	0.93
Derby et al. (1997)	Billy	1.00	1.00	1.00	2.14	1.00	0.93
	Danny	0.89	0.00	1.00	1.55	0.73	1.00
	Kelly	0.77	0.14	1.00	4.46	0.53	0.93
	Matt	0.94	0.85	0.96	1.16	0.89	0.94
Dunlap & Fox (1996)	Anthony	0.95	0.95	0.95	1.27	0.91	0.95
	Tom	1.00	1.00	1.00	1.80	1.03	1.00
	Larry	0.95	0.91	0.96	1.27	0.91	0.93
	Mario	1.00	1.00	1.00	2.08	1.02	1.00
	Michele	1.00	1.00	1.00	1.19	0.98	1.00
	Noah	1.00	1.00	1.00	1.60	0.95	1.00
Dunlap et al. (2006)	Alexis	0.91	0.89	0.89	3.95	0.90	0.94
	Maria	0.95	0.95	0.95	1.25	0.98	0.97
Harding et al. (2009a)	Andy	0.75	0.00	0.90	1.75	0.41	0.93
Harding et al. (2009b)	Al	1.00	1.00	1.00	8.43	1.07	1.00
	Kit	0.96	0.88	0.96	2.59	0.94	0.93
	Lou	0.89	0.80	0.87	1.63	0.79	0.89
Johnson et al. (2004)	Abe	0.95	0.77	1.00	1.79	0.94	0.91
	Child 1	0.86	0.47	1.00	1.74	0.84	0.63
	Child 2	1.00	1.00	1.00	2.93	1.00	1.03
	Child 3	0.98	0.90	1.00	0.82	0.93	0.96
Lucyshyn et al. (2007)	Katherine	0.95	0.89	0.94	2.16	0.93	0.94
Mancil et al. (2006)	Scott	1.00	1.00	1.00	2.77	1.00	0.98
Mancil et al. (2009)	David	0.97	0.87	1.00	2.06	0.92	1.02
	Scott	0.98	0.92	1.00	1.68	0.97	1.02
	Zeb	1.00	1.00	1.00	2.81	1.00	1.08

Olive et al. (2011)	Kerri	0.97	0.75	1.00	2.40	0.96	0.92
Padilla Dalmau et al. (2018)	Javier	0.67	0.00	0.73	1.07	0.69	0.14
Schieltz et al. (2011)	Sofia	0.96	0.75	1.00	2.13	0.93	1.07
	Bud	0.60	0.03	0.67	0.58	0.83	0.17
	Cam	0.69	0.17	0.86	1.35	0.80	0.31
	Juan	0.60	0.18	0.66	2.52	0.68	0.10
Schindler et al. (2005)	Ellie	0.89	0.64	0.98	1.76	0.90	0.77
	Kit	0.88	0.58	0.93	1.07	0.86	0.82
	Neal	0.94	0.83	0.94	1.62	0.93	0.90
Simacek et al. (2017)	Ella	0.76	0.18	0.91	1.06	0.79	0.48
	Lily	0.91	0.85	0.88	1.16	0.88	0.82
Wacker et al. (2011)	Jose	0.55	0.11	0.32	1.37	0.87	0.01
	Tina	0.77	0.02	0.86	1.35	0.85	0.49
Wacler et al. (2013a)	Jasper	0.87	0.09	0.91	1.51	0.88	0.75
	Kurt	0.96	0.80	1.00	3.02	0.92	0.84
	Rose	0.97	0.87	1.00	2.44	0.95	0.88
Wacker et al. (2013b)	Jack	0.74	0.43	0.81	0.97	0.88	0.55
	Jake	0.95	0.93	0.96	1.48	0.93	0.91
	Jill	0.61	0.17	0.75	0.77	0.89	0.21
	Mitt	0.65	0.08	0.92	0.93	0.89	0.31
	Tad	0.90	0.67	1.00	1.79	0.82	0.93
	Zeke	0.73	0.11	0.78	0.73	0.81	0.48

Appendix D
Effect Sizes for Communication

Study	Participant	NAP	PND	PEM	LRRU	Tau-U	PAND
Anderson et al. (2016)	Johnny	1.0	1.0	1.0	NA	1.0	1.0
Benson et al. (2017)	Nick	0.50	0.33	0.33	0.65	0.20	0.76
Berg et al. (2007)	Brent	0.87	0.74	0.99	2.33	0.76	0.93
	Calvin	1.0	1.0	1.0	6.71	0.98	1.0
Derby et al. (1997)	Theo	1.0	1.0	1.0	4.15	1.03	1.0
	Billy	0.97	0.97	0.97	NA	0.93	0.97
	Danny	1.0	1.0	1.0	5.97	1.0	1.0
	Kelly	1.0	1.0	1.00	NA	1.0	1.0
Harding et al. (2009a)	Matt	1.0	1.0	1.0	4.15	1.0	1.0
Harding et al. (2009b)	Andy	0.99	0.98	0.98	3.53	1.0	0.99
	Al	0.52	0	0.71	-0.13	0.07	0.82
	Kit	0.81	0	0.96	0.73	0.66	0.93
	Lou	0.93	0.87	0.93	1.93	0.65	0.91
	Abe	1.0	1.0	1.0	4.79	1.0	1.0
	David	1.0	1.0	1.00	6.55	1.0	1.0
Johnson et al. (2004)	Scott	1.0	1.0	1.0	5.28	1.0	1.0
Mancil et al. (2009)	Zeb	1.00	1.00	1.00	5.51	1.00	1.0
	Kerri	0.72	0	0.97	0.39	0.42	0.81
Olive et al. (2011)							
Padilla Dalmau et al. (2018)	Javier	1.0	1.0	1.0	5.33	1.0	1.0
	Sofia	1.0	1.0	1.0	6.3	0.96	1.0
Schindler et al. (2005)	Ellie	0.93	0.57	0.96	1.94	0.86	0.85
	Kit	0.87	0.69	0.86	4.59	0.76	0.85
	Neal	0.93	0.89	0.92	4.59	0.82	0.92
Simacek et al. (2017)	Ella	1.0	1.0	1.0	5.54	1.0	1.0
	Lily	0.79	0.51	0.85	1.34	0.58	0.85
Wacker et al. (2013a)	Jose	0.89	0.97	1.0	3.85	1.05	0.99
	Tina	0.88	0.76	0.98	3.27	0.71	0.98
	Jasper	0.41	0.26	0.27	-0.27	-0.18	0.83
	Kurt	0.81	0.35	0.93	1.48	0.61	0.90