Journal of
Applied Behavior Analysis

SEAB
SOCIETY FOR THE EXPERIMENTAL ANALYSIS OF BEHAVIOR

Journal of Applied Behavior Analysis 2020, **9999,** 1–22 NUMBER 9999 ()

Functional communication training: From efficacy to effectiveness

Mahshid Ghaemmaghami and Gregory P. Hanley

Department of Psychology, Western New England University and FTF Behavioral Consulting

Joshua Jessel

Department of Psychology, Western New England University and FTF Behavioral Consulting and Department of Psychology, Queens College

Functional communication training (FCT; Carr & Durand, 1985) is a common function-based treatment in which an alternative form of communication is taught to reduce problem behavior. FCT has been shown to result in substantial reductions of a variety of topographically and functionally different types of problem behavior in children and adults (efficacy). The extent to which these reductions maintain in relevant contexts and result in meaningful changes in the lives of those impacted (effectiveness) is the focus of this paper. This review evaluates the degree to which FCT has been established as an evidence-based practice in psychology (EBPP) according to the definition set out by the American Psychological Association's 2005 Presidential Task Force on Evidence-Based Practice. Our review finds overwhelming evidence in support of FCT as an efficacious treatment but highlights significant limitations in support of its effectiveness. In order to also be recognized as an EBPP, future research on FCT will need to focus more closely on issues related to home, school, and community application, feasibility, consumer satisfaction, and more general and global changes for the individual.

Key words: efficacy, effectiveness, functional communication training, generality, social validity

Problem behavior is an important public health concern due to its debilitating effects on the lives of the individuals exhibiting problem behavior and those around them. Behavioral science has led to the emergence of powerful technologies and procedures that have the potential to significantly improve the developmental and social trajectory of these individuals. For this potential to be realized, it is important to continuously monitor and reflect on the strength of evidence of routinely implemented behavioral technology. In general, a treatment literature begins with evaluations that determine whether a procedure can produce an effect. When effects are demonstrated and repeatedly replicated, evaluations progress to demonstrating that these effects can be maintained, transferred, and extended. The treatment literature on applied behavior analytic treatments for problem behavior has reached the point where it includes many demonstrations of positive, large effects, particularly with functional communication training (FCT), which is the most widely used and researched function-based behavioral treatment (Greer et al., 2016; Jessel et al., 2018; Kurtz et al., 2011; Rooker et al., 2013; Tiger et al., 2008). The current review is focused on examining the extent to which the literature on FCT has demonstrated these treatment effects can be maintained, transferred, and extended.

Functional Communication Training

FCT is a function-based treatment that involves teaching an alternative, and more

This review was conducted in partial fulfillment of a Ph.D. in Behavior Analysis from Western New England University by the first author.We thank Amanda Karsten, Jessica L. Sassi, Rachel Thompson, and Carolynn Kohn for their feedback on earlier versions of this manuscript.

[Correction added on 6 October 2020, after first online publication: The text citation for the reference 'Correa et al., 2019' has been corrected from 'Corea' to 'Correa'.]

Address correspondence to: Mahshid Ghaemmaghami, FTF Behavioral Consulting 40 Southbridge St., Suite 202, Worcester, MA 01608. Email: dr.g@ftfbc.com

doi: 10.1002/jaba.762

appropriate, functional communication response (FCR) that results in the delivery of the same reinforcer(s) purported to be maintaining problem behavior (Carr & Durand, 1985; Durand & Moskowitz, 2015). Tiger et al. (2008) identified three stages of FCT: (a) a functional assessment process including a functional analysis to identify the reinforcement contingency maintaining problem behavior, (b) the development of a socially acceptable FCR using the functional reinforcer, and (c) the extension of treatment to other settings and caregivers.

The success of FCT in reducing problem behavior coupled with its emphasis on teaching communication has made it a popular treatment for individuals with various intellectual and developmental disabilities. In the 35 years since the introduction of FCT by Carr and Durand (1985), FCT has been used successfully to treat various topographies of socially maintained problem behavior in adults and children (Durand & Moskowitz, 2015; Tiger et al. 2008). Functions treated using FCT include isolated contingencies (e.g., Fisher et al., 2000; Hagopian et al., 1998; Kurtz et al., 2003; Wacker et al., 2005) and synthesized contingencies (e.g., Bowman et al., 1997; Ghaemmaghami et al., 2016; Hagopian et al., 2007; Hanley et al., 2014; Sarno et al., 2011). The topographies treated include severe forms of problem behavior such as aggression, self-injury, and property destruction (e.g., Kurtz et al., 2011), and less severe forms such as autistic leading (Carr & Kemp, 1989), psychotic speech (Durand & Crimmins,1987), off-task behavior (Flood & Wilder, 2002), and ritualistic behavior (e.g., Rispoli et al., 2014). FCT has also been found suitable to be applied in conjunction with augmentative and alternative communication techniques (e.g., signs, communication boards; Mirenda, 1997).

The initial success of FCT appears to depend on a variety of factors (see Tiger et al., 2008, for a detailed review). These factors include the extent to which appropriate motivating operations are identified and contrived, the extent to which relevant reinforcers are provided contingent on the FCR instead of problem behavior, and the relative efficiency of the initial FCR. Extinction also appears to be a critical component of FCT (Fisher et al., 2000; Hagopian et al., 1998; Kurtz et al., 2011; Rooker et al., 2013). The inclusion of schedule thinning is another essential component of FCT (Durand & Moskowitz, 2015; Kurtz et al., 2011; Tiger et al., 2008). When delays to reinforcement are introduced, however, FCT with extinction often fails (Fisher et al., 2000; Hagopian et al., 1998; Hanley et al., 2001, Rooker et al., 2013) unless certain strategies are implemented as part of thinning reinforcement (see Hagopian et al., 2011, for a detailed review). Some of these strategies include establishing discriminative control of FCRs via multiple schedules (e.g., Greer et al., 2016; Hanley et al., 2001), chained schedules or demand fading (e.g., Falcomata et al., 2013; Lalli et al., 1995), and contingency-based delay tolerance training (e.g., Ghaemmaghami et al., 2016; Hanley et al., 2014; Jessel et al., 2018; Rose & Beaulieu, 2018).

Supplementary procedures have been added to FCT either before or during reinforcement thinning (Hagopian et al., 2011; Kurtz et al., 2011; Rooker et al., 2013). The FCT treatment algorithm provided by Kurtz et al. (2011) includes the addition of other treatment components to FCT as necessary. Punishment for problem behavior was previously the most common procedure added to FCT (Hagopian et al., 2011; Hagopian et al., 1998; Kurtz et al., 2011; Kurtz et al., 2003), but this has more recently shifted toward the use of additional reinforcement components (e.g., Greer et al., 2016; Hagopian et al., 2005; Kurtz et al., 2003; Rooker et al., 2013). Rooker et al. (2013) recently reported the addition of noncontingent reinforcement (NCR), differential reinforcement of

alternative behavior (DRA), or differential reinforcement of omission of problem behavior (DRO) to FCT to be more efficacious than the addition of punishment in their inpatient applications. Following the failure of FCT with extinction, the addition of alternative reinforcement resulted in a 90% reduction in 71% of application, whereas, the addition of punishment resulted in a 90% reduction in 54% of the applications. Hagopian et al. (2005) found that when thinning the schedule of reinforcement to practical levels, continuous and noncontingent access to stimuli that compete with the functional reinforcer is particularly useful in enhancing tolerance for long delays.

Efficacy versus Effectiveness of FCT

The terms *effective* and *efficacious* are often used interchangeably, but are, in fact, defined distinctly by the psychological and medical community (American Psychological Association, 2006). Treatment literature is often evaluated in terms of two attributes: *efficacy* and *effectiveness* (Glasgow et al., 2003; Hoagwood et al., 1995; Marchand et al., 2011; Smith et al., 2007). Efficacy relates to the demonstration of causal relations between the treatment in question and the change in behavior occurring under tightly controlled conditions. Effectiveness, on the other hand, relates to the clinical utility of treatment, its feasibility, generality, acceptability, and cost-effectiveness. In other words, efficacy is the potential of the treatment demonstrated under highly consistent conditions and strong experimental controls, whereas effectiveness is the actual effect of treatment in practice, where implementation conditions vary and multiple factors that may moderate the effects of the treatment are uncontrolled (Singal et al., 2014).

In general, there are some specific characteristics associated with studies examining efficacy versus effectiveness. Efficacy research tends to be conducted with relatively homogeneous populations, in highly controlled research settings, using single component or carefully described multicomponent treatments that are implemented with high integrity by skilled researchers for a short period of time (Singal et al., 2014). The outcome measures are often direct measures of the specific behavior treated, and the effects are demonstrated using strong experimental designs with high internal validity. Effectiveness research, by contrast, is the application of treatment with relatively heterogeneous populations, in typical settings where these treatments occur (e.g., schools, homes), using multicomponent treatments implemented with various ranges of integrity by caregivers, for typically much longer periods (Singal et al., 2014). The outcome measures include both direct and indirect assessments of (a) the specific target behavior; (b) social acceptability and adoptability of the intervention by the critical stakeholders; (c) cost-effectiveness; and (d) the more general impact of the intervention on adaptive functioning and the individual's overall quality of life.

The ultimate goal of applied research is to demonstrate that treatments derived from well-controlled efficacy studies remain useful in naturalistic studies at the level of clinical service, in order to establish the overall effectiveness of the intervention (Glasgow et al., 2003; Hoagwood et al. 1995; Marchand et al., 2011; Smith et al., 2007). For example, a treatment resulting in 90% reductions in tantrums or head-butting may be considered efficacious, but not effective if the change is not considered socially valid for that individual in his social environment (Wolf, 1978). Thus, quantified reductions in problem behavior identify the extent to which a treatment is efficacious, but are limited as a measure of the importance of the improvement which should be "large enough effects for practical value" (Baer et al., 1968, p. 96). Meeting such a goal requires valid research with emphasis on the social and ecological validity of the intervention, its procedures, and its effects.

The efficacy of FCT with extinction, under tightly controlled conditions and rich reinforcement schedules, has been demonstrated for a variety of topographically and functionally different problem behaviors, with a diverse range of participants varying in age, language and intellectual abilities, and comorbid diagnoses. Using the criteria set out by Division 12 of APA (Task Force, 1995) for empirically supported treatments (ESTs), Kurtz et al. (2011) found FCT with extinction far exceeds the criteria for a well-established treatment for socially maintained problem behavior of children with ID and ASD, and is probably efficacious for adults. Briefly, APA's EST criteria are largely concerned with the demonstrated efficacy of a treatment, as they consider the number of methodologically rigorous studies demonstrating an intervention is efficacious. Durand and Moskowitz (2015) further concluded that FCT exceeds the American Psychological Association's criteria for being designated a well-established treatment for the problem behavior of children with ID and other DDs including ASD. Despite the strong support for the efficacy of FCT, questions remain regarding its effectiveness.

Various models have been proposed for the evaluation of efficacy and effectiveness of psychological treatments. Some researchers advocate for phase models that view efficacy and effectiveness in a linear fashion and as distinct, and often opposite, phases of a complete evaluation. An intervention's efficacy is evaluated and confirmed before a question of effectiveness can be answered (Chambless & Hollon, 1998; Glasgow et al., 2003; Smith et al., 2007). Others hold a more flexible and bidirectional view of the relation between these constructs along a continuum that can be evaluated within the same study (Carroll & Rounsaville, 2003; Glasgow et al., 2006; Hallfors and Cho, 2007; Hoagwood et al., 2001; Hoagwood et al., 1995; Zwarenstein & Treweek, 2009; Zwarenstein et al., 2008). Inclusion of issues related to dissemination at the outset of the testing process may lessen the gap between research and practice caused by the overemphasis on efficacy concerns, resulting in treatments that are both efficacious and effective and ready for large-scale dissemination and consideration as an evidence-based practice in psychology (EBPP). This continuous and interactive view of efficacy and effectiveness is the model applied in this review.

The purpose of this review was to examine whether sufficient evidence exists with regard to the effectiveness of FCT such that we can establish FCT as an EBPP. We conducted a quantitative review of the literature to evaluate the strength of empirical evidence in support of the *effectiveness* of FCT in terms of the extent to which general (i.e., beyond the direct effects on the target behavior) and socially valid changes (i.e., change meaningful to relevant stakeholders) occur and maintain as a function of FCT in relevant natural contexts under manageable schedules of reinforcement. The specific questions investigated are the extent to which (a) large effects on target problem behavior have been obtained under rich and lean reinforcement schedules, with generality and long-term maintenance, and when implemented by caregivers in relevant contexts; (b) caregivers, behavior change agents and direct and indirect recipients of this treatment have socially validated the procedures and effects; and (c) secondary and more general effects of FCT have been evaluated and global improvements in functioning without severe side-effects have been demonstrated. The cost-effectiveness and feasibility of FCT and recommendations for future research are also discussed throughout.

Method

FCT studies were identified through a computer search of PsycINFO, PubMed, and Google Scholar using the keywords *functional*

communication training and *functional equivalence training* from 1985 to 2019. Additional articles were found through an examination of the reference lists of the identified FCT articles and reviews. The identified studies, and each individual FCT application within the study, were reviewed to determine those that met criteria for inclusion in this review.

Inclusion and Exclusion Criteria

Studies were included if they (a) were published in an English language peer-reviewed scholarly journal; (b) included an application of FCT as a component of treatment for one individual, topography of problem behavior, or behavior function; and (c) included graphic or aggregate pretreatment and posttreatment evaluation data on problem behavior for each individual application. Application refers to a single treatment evaluation of FCT for a participant's problem behavior from start to finish. In other words, all phases (e.g., schedule thinning phase, maintenance, treatment extension to other settings, tasks, or people) and comparisons (e.g. FCT with or without punishment, comparisons of mand modalities) used to identify the most successful treatment arrangement or combination were counted as one application. A new application of FCT was counted if different topographies or functions of problem behavior were treated separately. Relying on information provided by the authors and our own cross checking across articles, duplications of FCT applications across different studies were only counted once.

Coding Procedures

Studies that met the inclusion criteria were further evaluated to identify all applications of FCT. Each individual FCT application was analyzed for the following characteristics: (a) implementation context; (b) type of reinforcement schedule, (c) long-term maintenance, (d) generality, (e) social validation of procedures and effects, (f) positive secondary effects, (g) negative side effects and (h) general effects on functioning. The operational definitions of each characteristic coded are summarized in Table 1. For all items, data were also collected on whether positive effects were obtained. For items (a) to (d) positive effects were defined as large (i.e., 80% or more) reductions of baseline levels of problem behavior as reported by the authors (if available), or determined by using the average rate of combined problem behavior during the last three sessions of the first baseline and the last three sessions of the treatment phase in question. For the remaining items, positive effects were defined as either high levels of acceptability (e), any improvements from baseline (f and h), or absence of worsening (g).

Interrater Agreement

A second reader independently coded at least 35% of the articles and scored the FCT applications along the characteristics noted above. An item-by-item agreement was then calculated across the two application score sheets, with agreement being defined as both readers recording the same value for each characteristic. Interrater agreement averaged 98% (range, 89% to 100%) across applications.

Results and Discussion

A total of 208 empirical studies of FCT, published between 1985 and 2019, were included in this review. A total of 744 applications of FCT across 640 participants were identified. Figure 1 depicts the percentage of applications that show positive effects along the efficacy and effectiveness continuum. The size of each pie corresponds to the proportion of applications in each category out of 744 applications. Studies which were sourced for this review for analysis of data but were not individually cited in the body of the paper are listed in

Table 1

Operational Definitions of the Characteristics Coded for Each FCT Application

Coding Characteristic	Definitions
Expert Implementation	Experimenters and researchers, graduate, undergraduate students or therapists working as part of the research team, who were not the typical caregivers of the individual.
Caregiver Implementation	Parents, teachers, direct-care staff, and behavioral therapists and staff who regularly work with the individual, with or without experts present. Includes applications that started with caregiver implementation and those that were transferred to caregivers in relevant settings.
Analogue Settings	Tightly controlled environments such as in-patient and outpatient hospital units, session rooms in universities, segregated rooms in schools, and specialized clinics for assessment and treatment of problem behavior
Relevant Settings	Settings the individual is typically in and in which problem behavior was originally reported to occur.
Relevant Contexts	Situation in which problem behavior was originally reported to occur and which typically includes lean and unpredictable schedules of reinforcement
Rich SR Schedules	CRF schedule of reinforcement for the FCR following less than three demands or less than three seconds of delay.
Schedule Thinning	Procedures that thin the reinforcement schedule to any level beyond the CRF schedule noted above
Long-Term Maintenance	Treatment was implemented for an extensive period of six months or longer with continuous data collection or post-treatment follow-up probes. (Large effects were determined using the last phase).
Generality	Treatment was extended beyond the original teaching context, across tasks, people, and/or settings, with or without an analysis of generalization. (Large effects were determined across all secondary contexts evaluated).
Social Validation of Procedures	Acceptability of and satisfaction with overall procedures, or its components are measured either directly (e.g., concurrent chains arrangements) or indirectly (e.g., questionnaires, interviews) with corresponding data provided. This validation may have been provided by the client, direct caregivers, parents, or other stakeholders.
Social Validation of Effects	Acceptability and confirmation of meaningfulness of effects are measured either directly (e.g., scoring of sessions for severity and safety) or indirectly (e.g., questionnaires, interviews). This validation may have been provided by the client, direct caregivers, parents, or other stakeholders.
Positive Secondary Effects	Collateral effects of treatment increasing other adaptive behavior and/or reducing other non-target problem behavior, which were directly measured.
Negative Side-Effects	Collateral effects of treatment worsening existing adaptive behavior and/or increasing other non-target problem behavior, which were directly measured.
General Effects	Global effects on adaptive and play skills, IQ, language, diagnostic characteristics, or symptom severity are evaluated using standardized measures. General effects on family stress and functioning, community involvement, and overall quality of life are measured using either direct or indirect measures.

Note. FCT = Functional Communication Training, CRF = Continuous reinforcement, IQ = Intellectual Quotient

the Supplemental document (studies cited in the text are listed in the References).

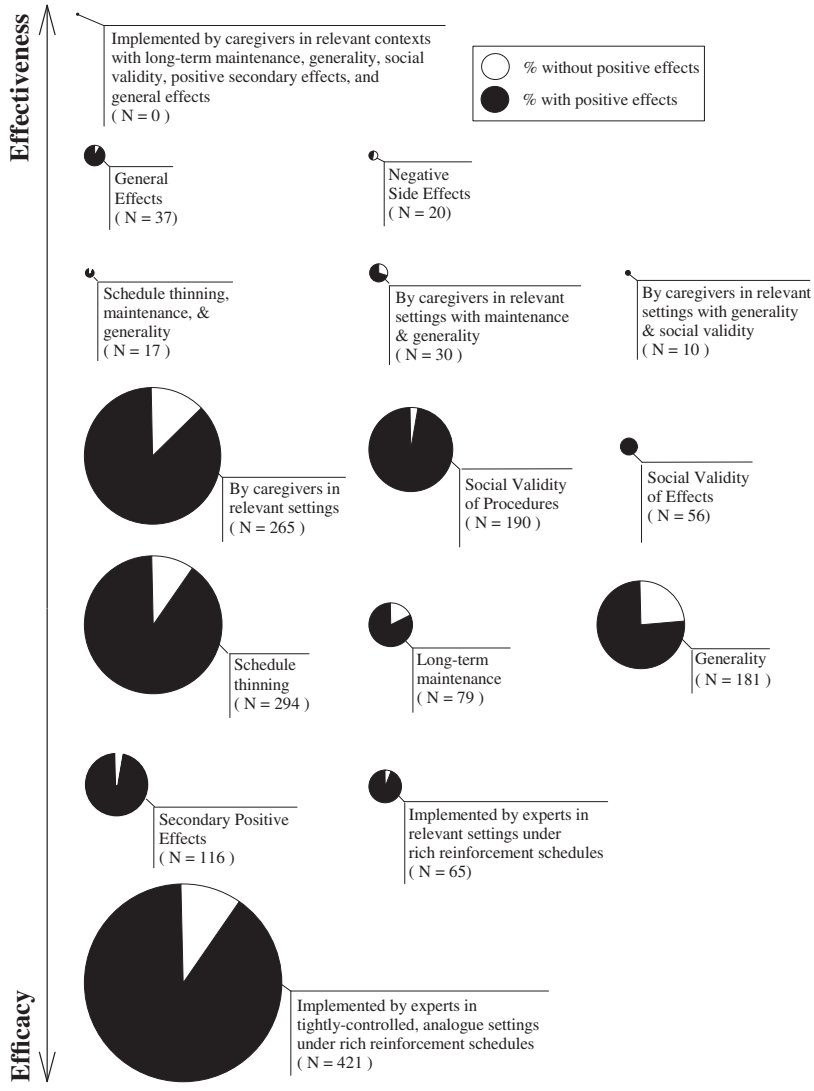
Large Effects on Target Problem Behavior

FCT results in a substantial reduction of problem behavior (reductions of 80% or more) in 90% of applications (668 out of 744) when it is first introduced. In addition, 57% (421 out of 744) of the applications were implemented by experts in highly controlled analogue settings under rich reinforcement contexts (Figure 1). Conversely, 9% (65 out of 744) of the applications were implemented by experts in relevant settings under rich reinforcement schedules and 94% of these have also shown large effects. Thus, there is strong evidence of efficacy, in line with previous reviews of FCT (Heath et al., 2015).

Schedule thinning, an essential component of FCT, was only included in 40% (294 out of 744) of the applications (Figure 1) which is similar to previous reviews. For example, Hagopian et al. (2011) found that only 19 out of 76 (i.e., 29%) studies published between 1985 and 2009 included a schedule thinning phase, and Kurtz et al. (2011) found that schedule thinning was only included in 32% of

Figure 1

Percentage of Functional Communication Training (FCT) Applications 1985-2019 With (Black Portion) and Without (White Portion) Large Positive Effects



Note. The size of the pie chart corresponds to the proportion of applications in each category out of the 744 applications.

the 106 applications. Large effects were obtained in 90% (265 out of 294) of the applications with schedule thinning (Figure 1). Most of the applications with schedule thinning, however, were implemented by experts (66%), and required the addition of supplemental procedures. This finding is similar to Kurtz et al. (2011) in which only 38% of successful applications with schedule thinning were done with FCT and extinction alone.

Some researchers have recognized the importance of identifying thinning procedures that increase the effectiveness of FCT under more typical reinforcement schedules and have

provided various recommendations for increasing the practicality of FCT (See Hagopian et al., 2011 for more details). Many researchers have used time-based arrangements using discriminative stimuli such as multiple schedules to incorporate periods of nonreinforcement of the communication response with or without response restriction (e.g., Fisher et al., 2014; Fisher et al., 1998; Greer et al., 2016; Greer et al., 2019; Hanley et al., 2001; Kuhn et al., 2010; Roane et al., 2004), others have used time-based mixed schedules such as delay schedules (e.g., Fisher et al., 2000; Hagopian et al., 1998; Rooker et al., 2013), demand fading in a chained schedule arrangement (e.g., Berg et al., 2007; Lalli et al., 1995; Wacker et al., 1998; Wacker, Lee, et al., 2013; Zangrillo et al., 2016), and contingency-based delay tolerance training (e.g., Carr & Carlson, 1993; Hanley et al., 2014; Jessel et al., 2018; Rose & Beaulieu 2018).

Despite the large reductions in problem behavior, complete elimination of problem behavior is difficult to achieve, especially under lean reinforcement schedules. Problem behavior was eliminated in 53% (398 out of 744) of the applications. For example, problem behavior was eliminated in 50% of the FCT applications in Wacker et al. (1998) and in less than 50% of the applications in Greer et al. (2016), Wacker et al. (2005), and Wacker, Lee, et al. (2013). All topographies of problem behavior were not eliminated in any applications by Kurtz et al. (2003) and self-injury was only eliminated when punishment was added to FCT. Rooker et al. (2013) found that problem behavior was eliminated in 19% of FCT applications with immediate reinforcement. Using discriminative stimuli (i.e., multiple schedules, response restriction, or chained schedules), Greer et al. found that problem behavior was eliminated in 40% of the applications with FCT alone.

In conclusion, FCT has resulted in large initial effects in 90% of the applications but eliminated problem behavior in only 53% of the applications. The majority of successful FCT applications (51%) have been implemented by experts in analogue settings under rich reinforcement schedules. Large effects under lean reinforcement have only been shown in 36% of FCT applications and usually only when supplemental procedures such as punishment (e.g., Fisher et al., 1993, Fisher et al., 2000) or competing reinforcement (e.g., Hagopian et al., 2005; Rooker et al., 2013) are added to FCT with extinction, or when multiple schedules (e.g., Fisher et al., 1998; Greer et al. 2016) or contingency-based delays (e.g., Ghaemmaghami et al., 2016; Jessel et al., 2018) or multicomponent interventions that include removal of the precipitating challenging situation via antecedent-based procedures (e.g., Carr et al., 1999; Kemp & Carr, 1995) are used.

In order to determine if sufficiently large effects on problem behavior are obtained using FCT, more research is needed. We do not yet have a reliable and systematic technology for determining the practical impact of an effect. Although, we can measure the degree to which a quantifiably large effect has been achieved, the extent to which it is a "large enough effect for practical value" (Baer et al., 1968, p. 96) is often not determined. One avenue to explore would be to examine the use of pre- and post-treatment questionnaires that identify the consumer's desired outcome, and their subsequent satisfaction with the outcomes along with the procedures. Descriptive information about the normative levels of problem behavior and functional repertoires in well-adjusted children and adults of various ages could help to establish a socially anchored success criterion.

More research is needed to determine when elimination of problem behavior should be the goal and the strategies that facilitate achieving this goal. Elimination of problem behavior is likely a goal for at least some recipients of FCT, but this is not achieved in many cases

when using immediate reinforcement of FCRs. Slaton et al. (2017) demonstrated that complete elimination of problem behavior may not be realized with FCT when co-occurring reinforcement contingencies are isolated, as some unaddressed establishing operation may result in the persistence of residual problem behavior. Synthesizing the co-occurring contingencies in the functional analysis and subsequent treatment resulted in immediate and complete elimination of problem behavior in all cases as compared to only 50% of cases when isolated contingencies were used. The relative advantages of synthesized contingencies when analyzing and treating problem behavior have been demonstrated several times subsequently (Mitteer et al., 2019; Slaton & Hanley, 2018; Zangrillo et al., 2016). More research on the relative effects of isolated versus synthesized contingencies on FCT outcomes is warranted, but reliance on isolated reinforcement contingencies should be reconsidered in light of these findings.

Additional research should also examine ways to consistently achieve the goal of elimination of problem behavior under leaned reinforcement schedules. Future studies might report separately the effects of FCT under rich versus lean schedules of reinforcement and across schedule thinning procedures. Because typical environments include sudden interruptions and unplanned denials of requests, procedures are needed for teaching tolerance for these events. The extent to which the various reinforcement thinning procedures are successful during unexpected interruptions remains to be determined.

One promising procedure designed explicitly for successful transfer to these natural contexts is the unpredictable contingency-based delay procedures described by Hanley et al. (2014) and replicated by Santiago et al. (2016). An intermittent and unpredictable reinforcement schedule is used. Some proportion of FCRs are reinforced immediately while others are followed by a brief delay signal (e.g., "Wait"). The learner is then taught to engage in progressively more complex and varied chains of contextually appropriate behavior before the reinforcer is delivered. Ghaemmaghami et al. (2016) showed that the response contingency during the delay interval was the factor driving the maintenance of near-zero rates of problem behavior and optimal rates of communication and other adaptive tolerance responses, especially as the delay intervals get longer. The extent to which the predictability and the variability of the response chain play a role in the success of this procedure remains to be evaluated.

Maintenance Over an Extensive Period of Time

Maintenance of effects can be evaluated as long-term implementation of FCT without a clear break between intervention and follow-up (e.g., Northup et al., 1994) or as posttreatment follow-up probes (e.g., Moes & Frea, 2002). Long-term maintenance was defined in this review as treatment being implemented for a period of 6 months or more, and was found in 11% (79 out of 744) of the applications (Figure 1). Only a small percentage of studies include data on the long-term maintenance of effects and only demand fading (e.g., Wacker et al., 2011; Wacker, Harding et al., 2013) and contingency-based delay tolerance training (e.g., Carr et al., 1999), have been evaluated for their long-term maintenance of effects.

The results for the long-term maintenance of effects are mixed but mostly show that reductions in problem behavior in the original training context maintain over time. Large effects were maintained in 82% (65 out of 79) of the applications (Figure 1). Although recovery to baseline levels of problem behavior are seldom reported, some problem behavior recovers over time, sometimes within a few weeks (e.g., Bailey et al., 2002).

Extensive long-term elimination, or near elimination, of problem behavior is often demonstrated in studies that combine FCT with other strategies in addition to extinction. For example, Carr et al. (1999) used an intervention package of FCT with choice-making, embedding demands, and contingency-based delays, and produced long-term maintenance of effects, including complete elimination of problem behavior for two out of three participants during the last 2 years of follow-up. Derby et al. (1997) combined FCT with punishment to achieve near elimination of problem behavior during the last year of follow-up. Moes & Frea (2002) used an FCT package that included parent training, inclusion of idiosyncratic family routines, extended family support, and full involvement of all family members to eliminate problem behavior. Maintenance of near-zero rates of problem behavior occurred for up to a year. The extent to which a treatment that includes only FCT, extinction, and a schedule thinning procedure without supplemental procedures would result in long term elimination of problem behavior remains largely unknown.

Future research on FCT should include examination of long-term maintenance as a standard part of the design. In addition, the specific variables that directly impact maintenance need to be identified. Lack of maintenance may reflect context renewal and resurgence (Bouton et al., 2012; Kelley et al., 2015), both of which are problematic when treatment integrity is diminished. Poor treatment integrity may be a result of low acceptability of procedures (e.g., Durand & Kishi, 1987; Steege et al., 1990;), poor recognizability of the FCR (Durand & Carr, 1991), or excessively high or diminished rates of the FCR (e.g., Northup et al., 1994). More research is needed to identify the roles of participant characteristics, family needs or strengths, caregiver values and concerns, and terminal schedules of reinforcement in the maintenance of effects with FCT.

Identifying variables that lead to better maintenance may also help researchers identify specific procedures for successful transfer of treatment effects to relevant settings. For example, Luczynski et al. (2014) found that generalization and maintenance of functional communication and self-control skills depended on informing children's teachers of the specific target skills and the importance of their intermittent reinforcement in the classroom. It is probable that maintenance of successful effects depends not only on teaching the individual the skills of communication and toleration but also alerting people in the individual's social context of the importance of reinforcement of these repertoires.

Transfer and Generalization to Other Contexts

Although FCT has the advantage of teaching behaviors that will access natural contingencies of reinforcement (Stokes & Baer, 1977), various studies have shown that generalization cannot be assumed (e.g., Horner & Budd, 1985; Olive et al., 2008; Rispoli et al., 2014; Schindler & Horner, 2005). Yet, evaluation of generality (i.e., treatment extension beyond the original teaching context) and generalization are not common in the FCT literature. Generality of treatment was evaluated in 24% (181 out of 744) of the applications (Figure 1); 76% of which showed large effects in secondary contexts. An analysis of generalization, however, has only been conducted in a handful of studies (e.g., Berg et al., 2007; Wacker et al., 2005).

The most commonly applied generalization tactic was *sequential modification* (Stokes & Baer, 1977) in which new tasks, people, or settings are successively introduced (e.g., Carr et al., 1999; Hanley et al., 2014; Rispoli et al., 2014). Other researchers have more directly programmed for generalization by incorporating *multiple exemplars* (Stokes &

Baer, 1977) through the use of multiple therapists and task materials from the beginning (e.g., Durand & Carr, 1991; Moes & Frea, 2002), *programming like stimuli* (Stokes & Baer, 1977) such as an augmentative speech device or a discriminative stimuli that can be transferred to new contexts (e.g., Durand, 1999; Olive et al., 2008), by initiating FCT in typical environments (e.g., Campbell & Lutzker, 1993; Santiago et al., 2016), or by incorporating familiar people or tasks from the typical environment into the training sessions (e.g., Hanley et al., 2014; Kemp & Carr, 1995). For example, Fisher et al. (2015) and Greer et al. (2019) have shown that the use of schedule-correlated stimuli under a multiple schedule arrangement may result in rapid transfer of direct effects of treatment to subsequent contexts with rich and lean reinforcement schedules.

The emergence of generalized responding is often observed following the addition of secondary or tertiary teaching materials (e.g., Mancil et al., 2009; Olive et al., 2008) or when a combination of generalization and social validity tactics are used (e.g., Durand, 1999; Durand & Carr, 1991; Hanley et al., 2014; Kemp & Carr, 1995). For example, Moes and Frea (2002) demonstrated generalization only following contextual modifications to FCT to incorporate idiosyncratic family situations such as caregiving demands, family support, and social interactions associated with each routine, as well as family values and goals, into the FCT procedures. Ghaemmaghami et al. (2016) showed that in addition to the use of multiple exemplars of antecedent conditions and delay cues, the type of reinforcement thinning procedure used may also impact generalization. Generalization occurred during contingency-based, but not during time-based, delay tolerance training.

Although not definitive, there is some evidence indicating generalization may be more frequently achieved across people and settings relative to tasks and activities (e.g., Berg et al., 2007; Wacker et al., 2005). Using multiple tasks and activities within multiple exemplar training appears important to increase the odds of generalization (e.g., Ghaemmaghami et al., 2016).

In conclusion, the generality of FCT has been evaluated in 24% of the applications, and large effects have been obtained in generality contexts in 76% of these cases. In addition, large effects have been obtained in 94% of the applications that include an evaluation of generality with long-term maintenance under lean schedules of reinforcement, but only 2% of applications include such an evaluation (Figure 1). Finally, these general effects have been obtained following the implementation of treatment in a tightly controlled environment and with deliberate programming (e.g., multiple exemplar training and sequential modification).

More research on efficient processes for achieving generality of the effects of FCT is needed. Although acquisition may be enhanced by relying on tightly controlled contexts, the extent to which this advantage is mitigated by difficulties with extending the effects of treatment outside that context should be examined. The demonstrations of generality that exist (e.g., Durand & Carr, 1992) have mostly been conducted under rich schedules. A few studies have evaluated generalization of FCT effects with various schedule thinning procedures (e.g., contingency-based and time-based delay, Ghaemmaghami et al., 2016; demand fading, Berg et al., 2007; Wacker et al., 2005; multiple schedules, Fisher et al., 2015; Greer et al., 2019), but more research is needed to clarify the type of reinforcement thinning procedures and supplemental procedures that should be added to FCT (e.g., schedule-correlated stimuli, punishment, tokens, timers), and whether relying on functional mediators (Stokes & Osnes, 1989) enhances generalization of FCT treatment effects. In particular,

the extent to which various reinforcement thinning procedures will result in less discriminable contingencies and the extent to which this increases generalization remains to be determined. Finally, assessments of the generalization of direct and indirect effects of FCT should be incorporated into the experimental designs of studies on FCT with comparisons of pretreatment and posttreatment probes. Problem behavior may not occur in all contexts nominated by caregivers, thus, the absence of problem behavior during posttreatment extension probes may reflect the absence of an establishing operation rather than a demonstration of generalization (Berg et al., 2007).

Effects of FCT with Caregivers in Relevant Settings

Caregiver implementation of FCT in relevant settings has been evaluated in 36% of the applications (265 out of 744) with 87% showing large effects (Figure 1). FCT has been implemented by parents, teachers, job coaches, and direct-care staff at homes, schools, vocational, and community settings. Some caregivers initially observe while the effects of treatment are evaluated by the research team and then implement the procedure in the natural environment with on-site coaching and feedback from the researchers (e.g., Campbell & Lutzker, 1993; Hanley et al., 2014; Kurtz et al., 2015). Others start with caregiver training and implementation in the natural environment (e.g., Carr & Carlson, 1993; Carr et al., 1999; Santiago et al., 2016; Wacker et al., 1998). More recently, Wacker, Lee et al. (2013) evaluated a training model in which parents initiated and delivered FCT treatments in a clinic with on-site support from a parent assistant and telehealth coaching by a behavior analyst. An average of 94% reduction in problem behavior of 17 children with ASD was obtained, although no data on FCRs and other collateral responses (e.g., task completion) were provided. Suess et al. (2014) also found acceptable levels of caregiver treatment fidelity following training provided via Telehealth.

The effects of FCT may be more quickly achieved when treatment is initially implemented by skilled researchers and then transferred to caregivers. For example, Derby et al. (1997) had all FCT and supplemental procedures initiated by parents in the natural environment. It took an average of 90 hr over 6 months (daily 10 to 30 min FCT session over 2 to 7 months) for the simple FCR to become 100% independent under immediate reinforcement. By contrast, Hanley et al. (2014) assessed and treated problem behavior incorporating delay tolerance training to practical demand levels, and treatment extension to parents and to the natural environment in an average of 27 hr distributed over 2 to 3 months.

Although 87% of caregiver-implemented applications of FCT in relevant contexts show large effects, problem behavior is not always eliminated. In addition, the effects are often obtained under rich reinforcement schedules in highly planned and specific teaching contexts without documentation of maintenance. Only 4% of FCT applications include an evaluation of generality and long-term maintenance when implemented by caregivers (Figure 1), and only 70% of these applications showed large effects. Studies that do show elimination of problem behavior under more naturalistic routines with caregivers are often multicomponent (e.g., Carr et al., 1999; Kemp & Carr 1995) rendering the importance of FCT unclear.

Although studies illustrate the efficacy of some parent training methods for delivering FCT (e.g., Wacker, Lee, et al., 2013; Suess et al., 2014), comparative studies of the efficacy and acceptability of the methods have not been conducted. Comparisons of parent training approaches might examine the speed and accuracy of the identification of functional reinforcers and teaching of communication skills,

the maintenance and generalization of effects, and the accuracy with which parents can arrange highly motivating evocative contexts. A cost–benefit analysis could be conducted for transferring implementation from the behavior analyst to the caregivers at the beginning of the process as opposed to the end and for a hybrid model (e.g., Campbell & Lutzker, 1993) in which caregivers are involved to some degree at every step. Finally, the effects of FCT with reinforcement thinning procedures as implemented by caregivers in relevant environments should be evaluated over a long period of time.

Social Validation of FCT Procedures and Effects

Wolf (1978) argued that issues of social validity are important from an ethical perspective and may be related to the long-term effectiveness of an intervention. The effects of treatment, generalization, and maintenance of FCT may be impacted by the acceptability of the procedures and effects (Moes & Frea, 2002; Northup et al., 1994; Steege et al., 1990). The majority of FCT studies, however, do not include any measures of social validity of the procedures, the effects, or both. Only 190 applications out of 744 (i.e., 26%) reported on social validity of the procedures with 97% reporting satisfaction. Only 56 (i.e., 8%) reported on the social validity of the effects (Figure 1) and 100% were satisfied. Finally, only 40 (i.e., 5%) reported on the validity of *both* procedures and effects and 100% were satisfied. The extent to which this may be an artifact of publication bias or "file drawer problem" remains unknown (Rosenthal, 1979).

To date, no researchers have identified the desired level of behavior change prior to the introduction of treatment. Additionally, the extent to which the degree of behavior change is meaningful and of practical value has only been evaluated in 8% (56 out of 744) of applications. Some questionnaires ask about hypothetical results (e.g., "Will this treatment be effective?"). Many questionnaires include questions about a change in the target behavior (e.g., "Did you observe behaviors improving as a result of the function-based treatment?" "How likely is treatment to improve behavior?"), but not whether the change was sufficient or of practical value. These social validity questions seem akin to indirect measures of the efficacy of FCT (i.e., there was a change) rather than measures of the social validity impact of the effects (i.e., the change was sufficient to result in a meaningful improvement in their quality of life).

A few studies have evaluated the acceptability of procedures and effects with caregivers (e.g., Hanley et al., 2014; Santiago et al., 2016; Jessel et al., 2018) and other stakeholders, such as vocational site employees and customers (e.g., Kemp & Carr, 1995), grocery store cashiers (e.g., Carr & Carlson, 1993), typical consumers such as parents (e.g., Dunlap et al., 2006), and group home staff (e.g., Carr et al., 1999). These studies include subjective assessments of the severity of the problem behavior exhibited by the participant in the relevant contexts and include questions pertaining to the comfort level and the willingness of the caregivers and stakeholders to interact with the participant following treatment.

Most studies include indirect measures of social validity in the form of questionnaires completed at the end of treatment. Some include measures of acceptability that are conducted before and after treatment (e.g., Hanley et al., 2014; Olive et al. 2008; Santiago et al., 2016). For example, Olive et al. (2008) asked the participant's parent to complete The Behavioral Intervention Rating Scale (BIRS; Elliott & Treuting, 1997) before and after FCT implementation to measure the level of acceptability of procedures and effects and to identify any changes in the consumer's attitude

pre- and postintervention. There is, however, a lack of repeated assessment of social validity. Most assessments of social validity occur once immediately after treatment but are rarely repeated throughout the treatment and at long term follow-up.

Participants themselves have sometimes directly validated the treatment they have received (Hanley et al., 1997; Hanley et al., 2005). Hanley et al. (1997) directly assessed children's preference for FCT, NCR, and extinction in a concurrent chains arrangement. Color-coded switches in the initial link were paired with 2 min of access to each treatment (FCT, NCR, extinction) in the terminal link. The cumulative initial link selections were used as a measure of preference for each treatment and children preferred FCT to both NCR and extinction. One study has directly evaluated caregiver preference for DRA treatments involving a communication response in place of problem behavior (i.e., FCT). Gabor et al. (2016) found that two out of five caregivers preferred DRA to DRO, NCR, or no treatment, two caregivers preferred differential reinforcement (either DRA or DRO) to NCR or no treatment, while one caregiver had no preference for any particular treatment but preferred these to no treatment.

In conclusion, despite the reported high acceptability of the procedures, there are some questions regarding the validity of the methods used to assess this acceptability (see Schwartz & Baer, 1991, for a more in-depth discussion of this issue), and limited evaluation of meaningfulness of the effects of these procedures. Thus, additional indirect and direct evaluations of the acceptability of the procedural variations of FCT are needed. In particular, researchers might examine participant and caregiver preferences for the various reinforcement thinning procedures. Although social validity questionnaires and interviews may overestimate consumer satisfaction (Schwartz & Baer, 1991) and may not be fully predictive of consumer behavior and long-term maintenance (Hawkins, 1991), these measures can nonetheless flag unsatisfactory procedures and effects. When detected, these findings provide an opportunity to address consumer concerns or to explore other approaches. Whether this may in turn enhance treatment implementation and outcome (Hawkins 1991; Schwartz & Baer, 1991) is an empirical question worthy of further investigation. In addition, researchers need more effective methods for identifying consumer preferences and satisfaction. Ultimately, procedures described by Welsh et al. (1994) in which continuous assessment, or survival probes, of maintained implementation of treatment by caregivers could be the measure used to determine acceptability of treatment.

Schwartz and Baer (1991), describe social validity assessment as a two-part process that includes using the information obtained from consumers to maintain and enhance the acceptability of treatment. Future research on FCT might examine the effects of a participatory action model (Correa et al., 2019; Fawcett, 1991; Ivankova, 2017; McCurdy et al., 2016; White, 2002) in which consumers of treatment are actively involved in the identification of treatment goads and outcomes and selection of treatment components. This model might be associated with higher treatment integrity as acceptability problems can be addressed at the beginning rather than discovered at the end. Thus far, only Moes and Frea (2002) have reported on how results of the social validity evaluations and parental values were incorporated into the selection of procedures and goals. Instead, social validity assessments have primarily been used to measure the extent to which caregivers approve of the procedures chosen by researchers and the effects of those procedures, which could lead to "false praise from consumers" (Schwartz & Baer, 1991, p. 191). Acceptability of treatment procedures without validation that the outcomes achieved are

sufficient and meaningful to the consumer, or vice versa is inadequate.

Acceptability of treatment may also be influenced by its overall feasibility which is directly impacted by its cost-effectiveness and requires an evaluation of the personnel, timeframe, and monetary costs associated with the effects obtained. Currently, this information is mostly absent from the literature on FCT. As we move toward an evaluation of FCT effectiveness and feasibility, future researchers should aim to provide the information necessary for a cost–benefit analysis (see Hanley et al., 2014 and Wacker, Lee, et al., 2013 for examples).

Effects of FCT on Global Functioning and Quality of Life

Freedom from a lifestyle dictated by problem behavior depends not just on the reduction of problem behavior via FCT but also on the extent to which restrictive lifestyles due to chemical and physical restraints, and seclusion and exclusion from community activities, and the negative impact on the overall family system are reduced or eliminated with this treatment (Fox et al., 2002). These secondary effects of treatment on functioning and quality of life are a critical part of the overall effectiveness of an intervention (Smith, 2012; Smith et al., 2007).

Secondary positive effects were evaluated in 15% (116 out of 744) of FCT applications, with 97% of these showing a positive effect on various adaptive responses or collateral reductions in other nontarget problem behavior (Figure 1). Adaptive responses included improved task engagement, play, social interaction, and spontaneous communication following FCT implementation. These positive effects are further increased when skill-based reinforcement thinning procedures are added to FCT, for example demand fading (e.g., Berg et al., 2007) or contingency-based delay tolerance training (e.g., Ghaemmaghami et al., 2016), which may be due to the explicit efforts to strengthen other contextually appropriate responses in addition to communication. Regarding collateral reductions, Wacker et al. (1998) reported a decrease in other minor topographies of problem behavior not directly treated by FCT and Scalzo et al. (2015) showed that FCT resulted in increased task engagement and lower levels of nontargeted challenging behavior. These effects may not be transient; Berg et al. (2007) showed the secondary effects of FCT on social interactions and task completion in both training and generalization contexts across settings and people.

Negative side effects of treatment were rarely evaluated. A total of 20 applications (out of 744) monitored negative side-effects, 11 (i.e., 55%) of which showed a worsening or a negative side-effect (Figure 1). For example, Fisher et al. (2000) showed that the introduction of delay resulted in an increase in stereotypy and inappropriate sexual behavior. Ghaemmaghami et al. (2016) reported on collateral effects such as excessive manding, emotional responding, and the emergence of novel topographies of problem behavior during time-based reinforcement delays. Similar to Fisher et al., time-based delays resulted in high levels of these negative collateral responses, whereas zero or near-zero rates of these responses were observed with contingency-based delays.

General effects of FCT on global functioning are also rarely evaluated. Only 5% of applications (37 out of 744) have evaluated general effects, with 92% of them showing a positive effect (Figure 1). Improvements in the quality of life following FCT implementation are rarely reported, except when FCT is part of a multicomponent intervention. For example, Carr & Carlson (1993) and Kemp & Carr (1995) noted significant improvements in their participants' ability to engage in community-based activities such as grocery shopping and vocational training. McConnachie and Carr (1997) found that FCT

resulted in lower levels of teacher self-reported stress, and more productive teaching sessions than escape extinction. Finally, Olive et al. (2008) found an increase in the overall rate of learning on standardized language tests, but no other researchers have examined general effects using global and standardized measures of adaptive and play skills, language, or intellectual quotient (IQ).

In summary, FCT has been shown to result in noteworthy increases in adaptive collateral responses such as compliance, play, task engagement, and social interaction, in addition to its direct effects on problem behavior. However, the extent to which these changes result in socially significant improvements in the overall adaptive functioning of the individual and overall improvements in quality of life, without adverse side-effects, remains largely unknown. Future research on FCT should include outcome measures of the indirect and general effects on adaptive functioning, quality of life, and the overall health of the individual and family. As reinforcement thinning procedures and other procedures are added to FCT, the relative utility of each combination may best be examined in their indirect effects on task engagement, compliance, social interaction, and emotional responding. Global measures such as the *Child Behavior Checklist* (Achenbach & Rescorla, 2000), the *Vineland Adaptive Behavior Scales, Third Edition* (Sparrow et al., 2016), the *Verbal Behavior Milestones Assessment and Placement Program* (Sundberg, 2008), or the *Peabody Picture Vocabulary Test – 4* (Dunn & Dunn, 2007) may be useful for detecting general effects of FCT in the type of large-scale randomized clinical trials that are valued by relevant stakeholders. The potential limitations of group designs (e.g., masking individual outcomes) can be addressed by combining group designs with single-subject analysis of individual data (e.g., Hagopian, 2020; Luczynski & Hanley, 2013).

Conclusions and Final Recommendations

This review indicates that implementation of FCT often results in large reductions in problem behavior while simultaneously increasing adaptive behaviors (e.g., communication, delay tolerance, social interaction). FCT is a flexible procedure that can be implemented by parents, teachers, and staff in many contexts. The procedure can be adapted to the needs of the individual (e.g., specific mode of communication) and the expectations of the social environment (e.g., completion of a work sequence prior to asking for a break).

Despite the breadth of these positive outcomes, the overwhelming majority of FCT research has focused on efficacy rather than effectiveness, researcher implementation rather than natural change agents, and rich reinforcement schedules rather than lean (bottom of Figure 1). There have been no applications of FCT implemented by caregivers in relevant contexts with demonstrated maintenance, generality, social validity, and positive secondary and general effects (top of Figure 1). This discrepancy between the amount of efficacy research and the amount of effectiveness research is also common for psychological therapies (Hoagwood et al., 1995). Nonetheless, both efficacy and effectiveness studies are important for a full exploration of the parameters of lasting and meaningful treatment benefits in clinical services. Currently the gap between these literatures requires a sizable inferential leap from highly controlled studies to family and practitioner implementation of FCT in typical settings.

The time seems right for a comprehensive evaluation of the effectiveness of FCT, and potentially other procedures. The most appropriate next step may be evaluations of efficacy and effectiveness within the same study incorporating the relevant experimental controls while also evaluating issues related to real world implementation. In the case of FCT, large scale, longitudinal, multicomponent evaluations that measure the feasibility, generality, and social validity of FCT will probably best fill the existing research-to-practice gap. The balance

of internal and ecological validity may be best achieved by starting in specialized or contrived settings with expert implementation and ending with caregivers implementing the procedures under practical reinforcement schedules and in relevant environments (as exemplified in Berg et al., 2007; Hanley et al., 2014). This type of research entails extensive long-term monitoring of treatment to capture more than the effects on problem behavior (as exemplified in Carr et al., 1999; Jensen et al., 2001) including the more general effects of treatment on individuals' overall functioning (as exemplified in Olive et al., 2008). Another design suited to addressing questions of effectiveness in behavior analytic research is the consecutive controlled case series (CCCS) which maintains the tenants of single-subject design while simultaneously addressing some questions of generality (see Hagopian, 2020; Hagopian et al., 2013 and Jessel et al., 2018). In a CCCS study, all cases that have received the treatment are included irrespective of outcome, which enables researchers to better identify any potential boundaries of generality and reduce the publication bias toward successful outcomes. The goal would be to establish the conditions for effectiveness of various procedural variations and guide practitioners' selection and implementation of treatment components based on variables such as participant characteristics, family values and sociocultural contexts.

In summary, there are significant strengths and limitations in the body of literature on FCT. The extent to which clinicians and families are encountering successful outcomes in general practice (i.e., effectiveness of FCT as an intervention) is not known because the effectiveness research is so limited. This research has likely not occurred because it is expensive and effortful for the researcher, but these barriers can be overcome. Both public and private grant funding could be allocated to effectiveness research that incorporates single- and between-subject designs that satisfy granting agencies. Another shift in contingencies could also be implemented through editorial prompts signaling reinforcement for including aspects of effectiveness (e.g., long-term evaluations, measures of social validity and secondary and general effects) in submitted research (e.g., Hanley, 2017).

FCT is an efficacious treatment that is ready to be evaluated as an evidence-based practice. In order to also be recognized as an effective treatment, as conveyed by Baer et al. (1968), and as an EBPP, future research on FCT will need to focus more closely on issues related to home, school, and community application, feasibility, and consumer satisfaction. What may be discovered is that FCT alone is not an EBPP, but rather, it is an essential component of a larger comprehensive treatment package that includes skill-based tolerance procedures that strengthen additional social skills such as compliance, task engagement, and social interaction, and therefore facilitate important global changes for the learner and his or her family members and teachers.

REFERENCES

Achenbach, T. M., & Rescorla, L. A. (2000). *Manual for the ASEBA preschool forms and profiles*. University of Vermont Department of Psychiatry.

American Psychological Association, Presidential Task Force on Evidence-Based Practice. (2006). Evidence-based practice in psychology. *American Psychologist*, *61*(4), 271–285. https://doi.org/10.1037/0003-066X.61.4.271.

Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis*, *1*(1), 91–97. https://doi.org/10.1901/jaba.1968.1-91.

Bailey, J., McComas, J. J., Benavides, C., & Lovascz, C. (2002). Functional assessment in a residential setting: Identifying an effective communicative replacement response for aggressive behavior. *Journal of Developmental and Physical Disabilities*, *14*(4), 353–369. https://doi.org/10.1023/A:1020382819146.

Berg, W. K., Wacker, D. P., Harding, J. W., Ganzer, J., & Barretto, A. (2007). An evaluation of multiple dependent variables across distinct classes of antecedent stimuli pre and post functional communication training. *Journal of Early and Intensive Behavior Intervention*, *4*(1), 305–333. https://doi.org/10.1037/h0100346.

Bouton, M. E., Winterbauer, N. E., & Todd, T. P. (2012). Relapse processes after the extinction of instrumental learning: Renewal, resurgence, and reacquisition. *Behavioural Processes*, *90*(1), 130–141. https://doi.org/10.1016/j.beproc.2012.03.004.

Bowman, L. G., Fisher, W. W., Thompson, R. H., & Piazza, C. C. (1997). On the relation of mands and the function of destructive behavior. *Journal of Applied Behavior Analysis*, *30*(2), 251–265. https://doi.org/10.1901/jaba.1997.30-251.

Campbell, R. V., & Lutzker, J. R. (1993). Using functional equivalence training to reduce severe challenging behavior: A case study. *Journal of Developmental and Physical Disabilities*, *5*(3), 203–216. https://doi.org/10.1007/BF01047064.

Carr, E. G., & Carlson, J. I. (1993). Reduction of severe behavior problems in the community using a multicomponent treatment approach. *Journal of Applied Behavior Analysis*, *26*(2), 157–172. https://doi.org/10.1901/jaba.1993.26-157.

Carr, E. G., & Durand, V. M. (1985). Reducing behavior problems through functional communication training. *Journal of Applied Behavior Analysis*, *18*(2), 111–126. https://doi.org/10.1901/jaba.1985.18-111.

Carr, E. G., & Kemp, D. C. (1989). Functional equivalence of autistic leading and communicative pointing: Analysis and treatment. *Journal of Autism and Developmental Disorders*, *19*(4), 561–578. https://doi.org/10.1007/BF02212858.

Carr, E. G., Levin, L., McConnachie, G., Carlson, J. I., Kemp, D. C., Smith, C. E., & McLaughlin, D. M. (1999). Comprehensive multisituational intervention for problem behavior in the community: Long-term maintenance and social validation. *Journal of Positive Behavior Interventions*, *1*(1), 5–25. https://doi.org/10.1177/109830079900100103.

Carroll, K. M., & Rounsaville, B. J. (2003). Bridging the gap: A hybrid model to link efficacy and effectiveness research in substance abuse treatment. *Psychiatric Services*, *54*(3), 333–339. https://doi.org/10.1176/appi.ps.54.3.333.

Chambless, D. L., & Hollon, S. D. (1998). Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology*, *66*(1), 7–18. https://doi.org/10.1037/0022-006X.66.1.7.

Correa, D. J., Kwon, C., Connors, S., Fureman, B., Whittemore, V., Jetté, N., … Moshé, S. L. (2019). Applying participatory action research in traumatic brain injury studies to prevent post-traumatic epilepsy. *Neurobiology of Disease*, *123*, 137–144. https://doi.org/10.1016/j.nbd.2018.07.007.

Derby, K. M., Wacker, D. P., Berg, W., Deraad, A., Ulrich, S., Asmus, J., … Stoner, E. A. (1997). The long-term effects of functional communication training in home settings. *Journal of Applied Behavior Analysis*, *30*(3), 507–531. https://doi.org/10.1901/jaba.1997.30-507.

Dunlap, G., Ester, T., Langhans, S., & Fox, L. (2006). Functional communication training with toddlers in home environments. *Journal of Early Intervention*, *28*(2), 81–96. https://doi.org/10.1177/105381510602800201.

Dunn, M., & Dunn, L. M. (2007). *Peabody Picture Vocabulary Test—4*. American Guidance Service, Inc.

Durand, V. M. (1999). Functional communication training using assistive devices: Recruiting natural communities of reinforcement. *Journal of Applied Behavior Analysis*, *32*(3), 247–267. https://doi.org/10.1901/jaba.1999.32-247.

Durand, V. M., & Carr, E. G. (1991). Functional communication training to reduce challenging behavior: Maintenance and application in new settings. *Journal of Applied Behavior Analysis*, *24*(2), 251–264. https://doi.org/10.1901/jaba.1991.24-251.

Durand, V. M., & Carr, E. G. (1992). An analysis of maintenance following functional communication training. *Journal of Applied Behavior Analysis*, *25*(4), 777–794. https://doi.org/10.1901/jaba.1992.25-777.

Durand, V. M., & Crimmins, D. B. (1987). Assessment and treatment of psychotic speech in an autistic child. *Journal of Autism and Developmental Disorders*, *17*(1), 17–28. https://doi.org/10.1007/BF01487257.

Durand, V. M., & Kishi, G. (1987). Reducing severe behavior problems among persons with dual sensory impairments: An evaluation of a technical assistance model. *Journal of the Association for Persons with Severe Handicaps*, *12*(1), 2–10. https://doi.org/10.1177/154079698701200102.

Durand, V. M., & Moskowitz, L. (2015). Functional communication training: Thirty years of treating challenging behavior. *Topics in Early Childhood Special Education*, *35*(2), 116–126. https://doi.org/10.1177/0271121415569509.

Elliott, S. N., & Treuting, M. V. B. (1997). The Behavior Intervention Rating Scale: Development and validation of a pretreatment acceptability and effectiveness measure. *Journal of School Psychology*, *29*(1), 43–51. https://doi.org/10.1016/0022-4405(91)90014-I.

Falcomata, T. S., Muething, C. S., Gainey, S., Hoffman, K., & Fragale, C. (2013). Further evaluations of functional communication training and chained schedules of reinforcement to treat multiple functions of challenging behavior. *Behavior Modification*, *37*(6), 723–746. https://doi.org/10.1177/0145445513500785.

Fawcett, S. B. (1991). Some values guiding community research and action. *Journal of Applied Behavior Analysis*, *24*(4), 612–636. https://doi.org/10.1901/jaba.1991.24-621.

Fisher, W. W., Greer, B. D., Fuhrman, A. M., & Querim, A. C. (2015). Using multiple schedules during functional communication training to promote rapid transfer of treatment effects. *Journal of Applied*

Behavior Analysis, *48*(4), 1–21. https://doi.org/10.1002/jaba.254.

Fisher, W. W., Greer, B. D., Querim, A. C., & DeRosa, N. (2014). Decreasing excessive functional communication responses while treating destructive behavior using response restriction. *Research in Developmental Disabilities*, *35*(11), 2614–2623. https://doi.org/10.1016/j.ridd.2014.06.024.

Fisher, W. W., Kuhn, D. E., & Thompson, R. H. (1998). Establishing discriminative control of responding using functional and alternative reinforcers during functional communication training. *Journal of Applied Behavior Analysis*, *31*(4), 543–560. https://doi.org/10.1901/jaba.1998.31-543.

Fisher, 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* (1), 23–36. https://doi.org/10.1901/jaba.1993.26-23.

Fisher, W. W., Thompson, R. H., Hagopian, L. P., Bowman, L. G., & Krug, A. (2000). Facilitating tolerance of delayed reinforcement during functional communication training. *Behavior Modification*, *24*(1), 3–29. https://doi.org/10.1177/0145445500241001.

Flood, W. A., & Wilder, D. A. (2002). Antecedent assessment and assessment-based treatment of off-task behavior in a child diagnosed with attention deficit-hyperactivity disorder (ADHD). *Education & Treatment of Children*, *25*(3), 331–338. https://www.jstor.org/stable/42899709.

Fox, L., Vaughn, B. J., Wyatte, M. L., & Dunlap, G. (2002). "We can't expect other people to understand": Family perspectives on problem behavior. *Exceptional Children*, *68*(4), 437–450. https://doi.org/10.1177/001440290206800402.

Gabor, A. M., Fritz, J. N., Roath, C. T., Rothe, B. R., & Gourley, D. A. (2016). Caregiver preference for reinforcement-based interventions for problem behavior maintained by positive reinforcement. *Journal of Applied Behavior Analysis*, *49*(2), 215–227. https://doi.org/10.1002/jaba.286.

Ghaemmaghami, M., Hanley, G. P., & Jessel, J. (2016). Contingencies promote delay tolerance. *Journal of Applied Behavior Analysis*, *49*(3), 548–575. https://doi.org/10.1002/jaba.333.

Glasgow, R., Green, L., Klesges, L., Abrams, D., Fisher, E., Goldstein, M., … Orleans, T. (2006). External validity: We need to do more. *Annals of Behavioral Medicine*, *31*(2), 105–108. https://doi.org/10.1207/s15324796abm3102_1.

Glasgow, R., Lichtenstein, E., & Marcus, A. (2003). Why don't we see more translation of health promotion research to practice? Rethinking the efficacy-to-effectiveness transition. *American Journal of Public Health*, *93*(8), 1261–1267. https://doi.org/10.2105/AJPH.93.8.1261.

Greer, B. D., Fisher, W. W., Briggs, A. M., Lichtblau, K. R., Phillips, L. A., & Mitteer, D. R. (2019). Using schedule-correlated stimuli during functional communication training to promote the rapid transfer of treatment effects. *Behavioral Development*, *24* (2), 100–119. https://doi.org/10.1037/bdb0000085.

Greer, B. D., Fisher, W. W., Saini, V., Owen, T. M., & Jones, J. K. (2016). Functional communication training during reinforcement schedule thinning: An analysis of 25 applications. *Journal of Applied Behavior Analysis*, *49* (1), 1–17. https://doi.org/10.1002/jaba.265.

Hagopian, L. P. (2020). The consecutive controlled case series: Design, data-analytics, and reporting methods supporting the study of generality. *Journal of Applied Behavior Analysis*, *53*(2), 596–619. https://doi.org/10.1002/jaba.691.

Hagopian, L. P., Boelter, E. W., & Jarmolowicz, D. P. (2011). Reinforcement schedule thinning following functional communication training: Review and recommendations. *Behavior Analysis in Practice*, *4*, 4–16. https://doi.org/10.1007/BF03391770.

Hagopian, L. P., Bruzek, J. L., Bowman, L. G., & Jennett, H. K. (2007). Assessment and treatment of problem behavior occasioned by interruption of free-operant behavior. *Journal of Applied Behavior Analysis*, *40*(1), 89–103. https://doi.org/10.1901/jaba.2007.63-05.

Hagopian, L. P., Contrucci Kuhn, S. A., Long, E. S., & Rush, K. S. (2005). Schedule thinning following communication training: Using competing stimuli to enhance tolerance to decrements in reinforcer density. *Journal of Applied Behavior Analysis*, *38*(2), 177–193. https://doi.org/10.1901/jaba.2005.43-04.

Hagopian, L. P., Fisher, W. W., Sullivan, M. T., Acquisto, J., & LeBlanc, L. A. (1998). Effectiveness of functional communication training with and without extinction and punishment: A summary of 21 inpatient cases. *Journal of Applied Behavior Analysis*, *31*(2), 211–235. https://doi.org/10.1901/jaba.1998.31-211.

Hagopian, L. P., Rooker, G. W., Jessel, J., & DeLeon, I. G. (2013). Initial functional analysis outcomes and modifications in pursuit of differentiation: A summary of 176 inpatient cases. *Journal of Applied Behavior Analysis*, *46*(1), 88–100. https://doi.org/10.1002/jaba.25.

Hallfors, D., & Cho, H. (2007). Moving behavioral science from efficacy to effectiveness. *International Journal of Behavioral Consultation and Therapy*, *3*(2), 236–250. https://doi.org/10.1037/h0100801.

Hanley, G. P. (2017). Editor's Note. *Journal of Applied Behavior Analysis*, *50*(1), 3–7. https://doi.org/10.1002/jaba.366.

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*(1), 17–38. https://doi.org/10.1901/jaba.2001.34-17.

Hanley, G. P., Jin, C. S., Vanselow, N. R., & Hanratty, L. A. (2014). Producing meaningful

improvements in problem behavior of children with autism via synthesized analyses and treatments. *Journal of Applied Behavior Analysis*, *47*(1), 16–36. https://doi.org/10.1002/jaba.106.

Hanley, G. P., Piazza, C. C., Fisher, W. W., Contrucci, S. A., & Maglieri, K. A. (1997). Evaluation of client preference for function-based treatment packages. *Journal of Applied Behavior Analysis*, *30*(3), 459–473. https://doi.org/10.1901/jaba.1997.30-459.

Hanley, G. P., Piazza, C. C., Fisher, W. W., & Maglieri, K. A. (2005). On the effectiveness of and preference for punishment and extinction components of function-based interventions. *Journal of Applied Behavior Analysis*, *38*(1), 51–65. https://doi.org/10.1901/jaba.2005.6-04.

Hawkins, R. P. (1991). Is social validity what we are interested in? Argument for a functional approach. *Journal of Applied Behavior Analysis*, *24*(2), 205–213. https://doi.org/10.1901/jaba.1991.24-205.

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*(2), 155–166. https://doi.org/10.1007/s40489-014-0044-3.

Hoagwood, K., Burns, B. J., Kiser, L., Ringeisen, H., & Schoenwald, S. K. (2001). Evidence-based practice in child and adolescent mental health services. *Psychiatric Services*, *52*(9), 1179–1189. https://doi.org/10.1176/appi.ps.52.9.1179.

Hoagwood, K., Hibbs, E., Brent, D., & Jensen, P. (1995). Introduction to the special section: Efficacy and effectiveness in studies of child and adolescent psychotherapy. *Journal of Consulting and Clinical Psychology*, *63*(5), 683–687. https://doi.org/10.1037/0022-006X.63.5.683.

Horner, R. H., & Budd, C. M. (1985). Acquisition of manual sign use: Collateral reduction of maladaptive behavior, and factors limiting generalization. *Education and Training of the Mentally Retarded*, *20*(1), 39–47. https://www.jstor.org/stable/23877281.

Ivankova, N. V. (2017). Applying mixed methods in community-based participatory action research: A framework for engaging stakeholders with research as a means for promoting patient-centredness. *Journal of Research in Nursing*, *22*(4), 282–294. https://doi.org/10.1177/1744987117699655.

Jensen, C. C., McConnachie, G., & Pierson, T. (2001). Long-term multicomponent intervention to reduce severe problem behavior: A 63-month evaluation. *Journal of Positive Behavior Interventions*, *3*(4), 225–236. https://doi.org/10.1177/109830070100300404.

Jessel, J., Ingvarsson, E. T., Metras, R., Kirk, H., & Whipple, R. (2018). Achieving socially significant reductions in problem behavior following the interview-informed synthesized contingency analysis: A summary of 25 outpatient applications. *Journal of Applied Behavior Analysis*, *51*(1), 130–157. https://doi.org/10.1002/jaba.436.

Kelley, M. E., Liddon, C. J., & Ribeiro, A. (2015). Basic and translational evaluation of renewal of operant responding. *Journal of Applied Behavior Analysis*, *48*(2), 390–401. https://doi.org/10.1002/jaba.209.

Kemp, D. C., & Carr, E. G. (1995). Reduction of severe problem behavior in community employment using an hypothesis-driven multicomponent intervention approach. *Journal of the Association for Persons with Severe Handicaps*, *20*(4), 229–247. https://doi.org/10.1177/154079699602000401.

Kuhn, D. E., Chirighin, A. E., & Zelenka, K. (2010). Discriminated functional communication: A procedural extension of functional communication training. *Journal of Applied Behavior Analysis*, *43*(2), 249–264. https://doi.org/10.1901/jaba.2010.43-249.

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*(6), 2935–2942. https://doi.org/10.1016/j.ridd.2011.05.009.

Kurtz, P. F., Chin, M. D., Huete, J. M., Tarbox, R. S., O'Connor, J. T., Paclawskyj, T. R., & Rush, K. S. (2003). Functional analysis and treatment of self-injurious behavior in young children: A summary of 30 cases. *Journal of Applied Behavior Analysis*, *36*(2), 205–219. https://doi.org/10.1901/jaba.2003.36-205.

Kurtz, P. F., Chin, M. D., Robinson, A. N., O'Connor, J. T., & Hagopian, L. P. (2015). Functional analysis and treatment of problem behavior exhibited by children with fragile X syndrome. *Research in Developmental Disabilities*, *43-44*, 150–166. https://doi.org/10.1016/j.ridd.2015.06.010.

Lalli, J. S., Casey, S., & Kates, K. (1995). Reducing escape behavior and increasing task completion with functional communication training, extinction, and response chaining. *Journal of Applied Behavior Analysis*, *28*(3), 261–268. https://doi.org/10.1901/jaba.1995.28-261.

Luczynski, K. C., & Hanley, G. P. (2013). Preventing the development of problem behavior by teaching functional communication and self-control skills to preschoolers. *Journal of Applied Behavior Analysis*, *46*(2), 355–368. https://doi.org/10.1002/jaba.44.

Luczynski, K. C., Hanley, G. P., & Rodriguez, N. M. (2014). An evaluation of the generalization and maintenance of functional communication and self-control skills with preschoolers. *Journal of Applied Behavior Analysis*, *47*(2), 246–263. https://doi.org/10.1002/jaba.128.

Mancil, G. R., Conroy, M. A., & Haydon, T. F. (2009). Effects of a modified milieu therapy intervention on the social communicative behaviors of young children

with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *39*(1), 149–163. https://doi.org/10.1007/s10803-008-0613-3.

Marchand, E., Stice, E., Rohde, P., & Becker, C. B. (2011). Moving from efficacy to effectiveness trials in prevention research. *Behaviour Research and Therapy*, *49*(1), 32–41. https://doi.org/10.1016/j.brat.2010.10.008.

McConnachie, G., & Carr, E. G. (1997). The effects of child behavior problems on the maintenance of intervention fidelity. *Behavior Modification*, *21*(2), 123–158. https://doi.org/10.1177/01454455970212001.

McCurdy, B. L., Thomas, L., Truckenmiller, A., Rich, S. H., Hillis-Clark, P., & Lopez, J. C. (2016). School-wide positive behavioral interventions and supports for students with emotional and behavioral disorders. *Psychology in the Schools*, *53*(4), 375–389. https://doi.org/10.1002/pits.21913.

Mirenda, P. (1997). Supporting individuals with challenging behavior through functional communication training and AAC: Research review. *Augmentative and Alternative Communication*, *13*(4), 207–225. https://doi.org/10.1080/07434619712331278048.

Mitteer, D. R., Fisher, W. W., Briggs, A. M., Greer, B. D., & Hardee, A. M. (2019). Evaluation of an omnibus mand in the treatment of multiply controlled destructive behavior. *Behavioral Development*, *24*(2), 74–88. https://doi.org/10.1037/bdb0000088.

Moes, D. R., & Frea, W. D. (2002). Contextualized behavioral support in early intervention for children with autism and their families. *Journal of Autism and Developmental Disorders*, *32*(6), 519–533. https://doi.org/10.1023/A:1021298729297.

Northup, J., Wacker, D. P., Berg, W. K., Kelly, L., Sasso, G., & DeRaad, A. (1994). The treatment of severe behavior problems in school settings using a technical assistance model. *Journal of Applied Behavior Analysis*, *27*(1), 33–47. https://doi.org/10.1901/jaba.1994.27-33.

Olive, M. L., Lang, R. B., & Davis, T. N. (2008). An analysis of the effects of functional communication and a voice output communication aid for a child with autism spectrum disorder. *Research in Autism Spectrum Disorders*, *2*(2), 223–236. https://doi.org/10.1016/j.rasd.2007.06.002.

Rispoli, M., Camargo, S., Machalicek, W., Lang, R., & Sigafoos, J. (2014). Functional communication training in the treatment of problem behavior maintained by access to rituals. *Journal of Applied Behavior Analysis*, *47*(3), 580–593. https://doi.org/10.1002/jaba.130.

Roane, H. S., Fisher, W. W., Sgro, G. M., Falcomata, T. S., & Pabico, R. R. (2004). An alternative method of thinning reinforcer delivery during differential reinforcement. *Journal of Applied Behavior Analysis*, *37*(2), 213–218. https://doi.org/10.1901/jaba.2004.37-213.

Rooker, G. W., Jessel, J., Kurtz, P. F., & Hagopian, L. P. (2013). Functional communication training with and without alternative reinforcement and punishment: An analysis of 58 applications. *Journal of Applied Behavior Analysis*, *46*(4), 708–722. https://doi.org/10.1002/jaba.76.

Rose, J. C., & Beaulieu, L. (2018). Assessing the generality and durability of interview-informed functional analyses and treatment. *Journal of Applied Behavior Analysis*, *52*(1), 271–285. https://doi.org/10.1002/jaba.504.

Rosenthal, R. (1979). File drawer problem and tolerance for null results. *Psychological Bulletin*, *86*(3), 638–641. https://doi.org/10.1037/0033-2909.86.3.638.

Santiago, J. L., Hanley, G. P., Moore, K., & Jin, C. S. (2016). The generality of interview-informed functional analyses: Systematic replication in school and home. *Journal of Autism and Developmental Disorders*, *46*(3), 797–811. https://doi.org/10.1007/s10803-015-2617-0.

Sarno, J. M., Sterling, H. E., Mueller, M. M., Dufrene, B., Tingstrom, D. H., & Olmi, D. J. (2011). Escape-to-attention as a potential variable for maintaining problem behavior in the school setting. *School Psychology Review*, *40*(1), 57-71. https://www.tandfonline.com/doi/abs/10.1080/02796015.2011.12087728

Scalzo, R., Henry, K., Davis, T. N., Amos, K., Zoch, T., Turchan, S., & Wagner, T. (2015). Evaluation of interventions to reduce multiply controlled vocal stereotypy. *Behavior Modification*, *39*(4), 496–509. https://doi.org/10.1177/0145445515573986.

Schindler, H. R., & Horner, R. H. (2005). Generalized reduction of problem behavior of young children with autism: Building trans-situational interventions. *American Journal on Mental Retardation*, *110*(1), 36-47. https://doi.org/10.1352/0895-8017(2005)110<36:GROPBO>2.0.CO;2

Schwartz, I. S., & Baer, D. M. (1991). Social validity assessments: Is current practice state of the art? *Journal of Applied Behavior Analysis*, *24*(2), 189–204. https://doi.org/10.1901/jaba.1991.24-189.

Singal, A. G., Higgins, P. D., & Waljee, A. K. (2014). A primer on effectiveness and efficacy trials. *Clinical and Translational Gastroenterology*, *5*(1), e45. https://doi.org/10.1038/ctg.2013.13.

Slaton, J. D., & Hanley, G. P. (2018). Nature and scope of synthesis in functional analysis of problem behavior. *Journal of Applied Behavior Analysis*, *51*(4), 943–973. https://doi.org/10.1002/jaba.498.

Slaton, J. D., Hanley, G. P., & Raftery, K. (2017). Interview-informed functional analyses: A comparison of synthesized and isolated variables. *Journal of Applied Behavior Analysis*, *50*(2), 252–277. https://doi.org/10.1002/jaba.384.

Smith, T. (2012). Evolution of research on interventions for individuals with autism spectrum disorder:

Implications for behavior analysts. *The Behavior Analyst*, *35*(1), 101–113. https://doi.org/10.1007/BF03392269.

Smith, T., Scahill, L., Dawson, G., Guthrie, D., Lord, C., Odom, S., … Wagner, A. (2007). Designing research studies on psychosocial interventions in autism. *Journal of Autism and Developmental Disorders*, *37*(2), 354–366. https://doi.org/10.1007/s10803-006-0173-3.

Sparrow, S. S., Cicchetti, D. V., & Saulnier, C. A. (2016). *Vineland Adaptive Behavior Scales*, *Third Edition (Vineland-3)*. Pearson.

Steege, M. W., Wacker, D. P., Cigrand, K. C., Berg, W. K., Novak, C. G., Reimers, T. M., … DeRaad, A. (1990). Use of negative reinforcement in the treatment of self-injurious behavior. *Journal of Applied Behavior Analysis*, *23*(4), 459–467. https://doi.org/10.1901/jaba.1990.23-459.

Stokes, T. F., & Baer, D. M. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis*, *10*(2), 349–367. https://doi.org/10.1901/jaba.1977.10-349.

Stokes, T. F., & Osnes, P. G. (1989). An operant pursuit of generalization. *Behavior Therapy*, *20*(3), 337–355. https://doi.org/10.1016/S0005-7894(89)80054-1.

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*(1), 34–59. https://doi.org/10.1007/s10864-013-9183-3.

Sundberg, M. L. (2008). *Verbal Behavior Milestones Assessment and Placement Program (VB-MAPP)*. AVB Press.

Task Force on Promotion and Dissemination of Psychological Procedures. (1995). Training in and dissemination of empirically-validated psychological treatments. *The Clinical Psychologist*, *48*(1), 3–23.

Tiger, J. H., Hanley, G. P., & Bruzek, J. (2008). Functional communication training: A review and practical guide. *Behavior Analysis in Practice*, *1*(1), 16–23. https://doi.org/10.1007/BF03391716.

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*(2-3), 233–256. https://doi.org/10.1080/0144341042000301184.

Wacker, D. P., Berg, W. K., Harding, J. W., Derby, M. K., Asmus, J. M., & Healy, A. (1998). Evaluation and long-term treatment of aberrant behavior displayed by young children with disabilities. *Journal of Developmental & Behavioral Pediatrics*, *19*(4), 260–266. https://insights.ovid.com/article/00004703-199808000-00004.

Wacker, D. P., Harding, J. W., Berg, W. K., Lee, J. F., Schieltz, K. M., Padilla, Y. C., … Shahan, T. A. (2011). An evaluation of persistence of treatment effects during long-term treatment of destructive behavior. *Journal of the Experimental Analysis of Behavior*, *96*(2), 261–282. https://doi.org/10.1901/jeab.2011.96-261.

Wacker, D. P., Harding, J. W., Morgan, T. A., Berg, W. K., Schieltz, K. M., Lee, J. F., & Padilla, Y. C. (2013). An evaluation of resurgence during functional communication training. *The Psychological Record*, *63*(1), 3–20. https://doi.org/10.11133/j.tpr.2013.63.1.001.

Wacker, D. P., Lee, J. F., Dalmau, Y. C. . P., Kopelman, T. G., Lindgren, S. D., Kuhle, J., … Waldron, D. B. (2013). Conducting functional communication training via telehealth to reduce the problem behavior of young children with autism. *Journal of Developmental and Physical Disabilities*, *25*(1), 35–48. https://doi.org/10.1007/s10882-012-9314-0.

Welsh, T. M., Miller, L. K., & Altus, D. E. (1994). Programming for survival: A meeting system that survives 8 years later. *Journal of Applied Behavior Analysis*, *27*(3), 423–433. https://doi.org/10.1901/jaba.1994.27-423.

White, G. W. (2002). Consumer participation in disability research: The golden rule as a guide for ethical practice. *Rehabilitation Psychology*, *47*(4), 438–446. https://doi.org/10.1037/0090-5550.47.4.438.

Wolf, M. M. (1978). Social validity: The case for subjective measurement or how applied behavior analysis is finding its heart. *Journal of Applied Behavior Analysis*, *11*(2), 203–214. https://doi.org/10.1901/jaba.1978.11-203.

Zangrillo, A. N., Fisher, W. W., Greer, B. D., Owen, T. M., & DeSouza, A. A. (2016). Treatment of escape-maintained challenging behavior using chained schedules: An evaluation of the effects of thinning positive plus negative reinforcement during functional communication training. *International Journal of Development Disabilities*, *62*(3), 147–156. https://doi.org/10.1080/20473869.2016.1176308.

Zwarentstein, M., & Treweek, S. (2009). What kind of randomized trials do we need? *Canadian Medical Association Journal*, *180*(10), 998–1000. https://doi.org/10.1503/cmaj.082007.

Zwarenstein, M., Treweek, S., Gagnier, J., Altman, D., Tunis, S., Haynes, B., … Moher, D. (2008). Improving the reporting of pragmatic trials: An extension of the CONSORT statement. *British Medical Journal*, *337*, a2390–a2397. https://doi.org/10.1136/bmj.a2390.

Received February 2, 2018
Final acceptance July 23, 2020
Action Editor, Louis Hagopian

Supporting information

Additional Supporting Information may be found in the online version of this article at the publisher's website.