

# Contingency-based delay to reinforcement following functional communication training for autistic individuals: A multilevel meta-analysis

Autism  
2022, Vol. 26(4) 761–781  
© The Author(s) 2021  
Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/13623613211065540  
journals.sagepub.com/home/aut  


Reem Muharib<sup>1</sup>, Art Dowdy<sup>2</sup> , Adithyan Rajaraman<sup>3</sup>  
and Joshua Jessel<sup>4</sup> 

## Abstract

Functional communication training, an intervention for challenging behavior rooted in principles of applied behavior analysis, has copious empirical support dating back to the mid-1980s for autistic individuals. Recently, there has been a concerted effort to thin reinforcement delivery during functional communication training using contingency-based delays that, in turn, are designed to enhance practicality and feasibility while not compromising on efficacy. In this synthesis, we meta-analyzed the literature using log response ratio effect sizes to investigate (a) combined and across type effectiveness of contingency-based delays and (b) moderating variables that might impact intervention outcomes. Findings showed that contingency-based delays were effective for autistic individuals (log response ratio = -2.17; 95% CI = (-2.76, -1.58)) and most effective when the contingency incorporated positive reinforcement (log response ratio = -2.30; 95% CI = (-2.83, -1.78)). In addition, delay procedures that included differential reinforcement of alternative behavior were overall more effective (log response ratio = -2.13; 95% CI = (-2.72, -1.55)) than those that involved differential reinforcement of other behavior (log response ratio = -1.24; 95% CI = (-3.84, 1.37)). Noteworthy moderating variables found to impact contingency-based delay efficacy included the intervention dosage and the topography of behavior. We discuss these findings and highlight directions where additional empirical research is warranted to improve our understanding about contingency-based delays for autistic individuals.

## Lay abstract

Functional communication training, an intervention for challenging behavior rooted in principles of applied behavior analysis, has copious empirical support dating back to the mid-1980s for autistic individuals. Recently, there has been a concerted effort to thin reinforcement delivery during functional communication training using contingency-based delays that, in turn, are designed to enhance practicality and feasibility while not compromising efficacy. In this synthesis, we meta-analyzed the literature base with the goal of investigating both combined and across type effectiveness of contingency-based delays. We also aimed to investigate moderating variables that might impact intervention outcomes. Findings showed that contingency-based delays were effective for individuals with an autism spectrum disorder diagnosis and most effective when the delay incorporated some form of positive reinforcement. In addition, differential reinforcement of alternative-based delays was overall more effective when compared to differential reinforcement of other behavior-based delays. Noteworthy moderating variables found to impact contingency-based delay efficacy included the intervention dosage and the topography of behavior. We discuss these findings and highlight directions where additional empirical research is warranted to improve our understanding about contingency-based delays for individuals diagnosed with autism spectrum disorder.

<sup>1</sup>Texas State University, USA

<sup>2</sup>Temple University, USA

<sup>3</sup>University of Maryland, Baltimore County (UMBC), USA

<sup>4</sup>Queens College, USA

## Corresponding author:

Art Dowdy, College of Education and Human Development, Temple University, 1301 Cecil B. Moore Ave., Philadelphia, PA 19122, USA.  
Email: dowdy@temple.edu

## Keywords

applied behavior analysis, autism spectrum disorder, challenging behavior, log response ratio, multilevel meta-analysis, reinforcement schedule thinning

Autistic individuals sometimes develop intractable challenging behavior such as aggression (Kanne & Mazurek, 2011) and self-injurious behavior (SIB; Richards et al., 2012). Although challenging behavior is not a required defining characteristic to be formally diagnosed with autism spectrum disorder (ASD), autistic individuals are at a greater risk than their typically developing peers to exhibit challenging behavior (Adamek et al., 2011). In fact, autistic individuals have a high rate of comorbid behavioral, mood, and anxiety disorders (Leyfer et al., 2006; Matson & Nebel-Schwalm, 2007; Simonoff et al., 2008), symptoms of which all include overt expressions of challenging behavior.

Challenging behavior has been shown to adversely impact educational achievement and community participation, increase the risk of hospitalization and admission to residential care (Mandell, 2008), and increase parental stress (Schiltz et al., 2018) that, in turn, could manifest as potential criticism directed at autistic individuals (Baker et al., 2019). Chiang and Wineman (2014) found that without the introduction of a formal intervention, challenging behavior is likely to be an ongoing and chronic concern that adversely impacts an autistic person's quality of life (QoL). Behavioral intervention has been shown to support the dynamic needs of autistic individuals and is particularly critical upon the first signs of challenging behavior (Reichow, 2012; Ruppel et al., 2021; Tincani et al., 2018). Early behavioral intervention lessens the likelihood of unmanageable challenging behavior that can result in costly inpatient services and increases the likelihood of living an independent lifestyle with minimal supports (Lowe et al., 2007; Mandell, 2008).

Essential components of any behavioral intervention often involve multiple strategies focusing on (a) modifying antecedent events to proactively support alternatives to challenging behavior before they occur, (b) arranging consequences to reactively strengthen alternatives to challenging behavior after they occur, and (c) targeting developmentally appropriate and sustainable skills to replace challenging behavior. While behavioral interventions can include an eclectic collection of any number of these strategies, evidence-based practices aimed at a comprehensive level of support designed to teach adaptive, behavioral, and academic skills have been developed to meet the complex needs of autistic individuals (Hume et al., 2021).

Programming a comprehensive behavioral intervention to reduce challenging behavior often involves three steps (Hanley et al., 2014; Santiago et al., 2016). First, the clinician conducts a functional assessment to determine the

context in which challenging behavior is likely to occur. The functional assessment allows the clinician to identify environmental variables that are contributing to the occurrence of challenging behavior. This can include information on negative reinforcers contributing to challenging behavior (e.g. avoiding work requirements), positive reinforcers contributing to challenging behavior (e.g. access to preferred activities), or more likely than not, some combination of reinforcers contributing to challenging behavior (e.g. avoiding work to access preferred activities).

With the information obtained from the results of the functional assessment, the clinician can begin the second step: implementing a skill-based intervention. The skill-based intervention focuses on strengthening deficits in communication skills and has been termed functional communication training (FCT; Carr & Durand, 1985). During FCT the clinician provides the functional<sup>1</sup> reinforcers contingent on progressively more complex language abilities before reaching a socially and developmentally acceptable targeted goal (Ghaemmghami et al., 2018). By the final step, extensions are added to intervention procedures to improve generality and maintenance of reductions in challenging behavior in ecologically relevant situations. This often includes training caregivers on intervention implementation, teaching and implementing intervention in a variety of settings, and reinforcement schedule thinning (RST; Ghaemmghami et al., 2021).

RST, which involves progressively extending periods of time in which reinforcement is unavailable, has garnered recent interest as an essential extension for clinicians to implement during behavioral intervention for challenging behavior (Hagopian et al., 2011). This is particularly important considering that most often, initial steps of a behavioral intervention involve much effort on professionals' and caretakers' behalf to ensure that reinforcement is provided frequently, immediately, and often after each occurrence of the newly taught behavior. For example, in initial FCT, the clinician typically provides the functional reinforcer following each and every communicative response emitted by the individual. Although a crucial step to rapidly reduce the occurrence of challenging behavior, the effort required of clinicians and caretakers to sustain this level of intervention is often impractical and not typical of the home or school environment. In the home environment, caregiver expectations may not only be for the child to appropriately request access to preferred items but to accept denials when those items are not available, such as when it is a sibling's turn. By contrast, teachers will often need their students to tolerate periods of structured activities throughout the day. That is to say, FCT alone has

often been reported as successful, evidence-based practice among experts in tightly controlled settings, but evidence of the generality of FCT among more socially relevant change agents in the natural environment remains to be fully explored (Ghaemmaghmi et al., 2021). In fact, surveys and interviews of parents' experiences with professionals found that formal support services often resulted in families feeling isolated and alienated, thus lacking a family centered approach to care (Galpin et al., 2018). When providing behavioral intervention, RST is, therefore, a critical step to ensure that intervention matches the needs specific to caretakers, educational personnel, and most importantly autistic individuals.

Multiple procedural variations of RST exist (Muharib et al., 2019) and can be grouped under two general categories, which are distinguished by the programmed criteria for the return of the reinforcers following the scheduled delay (Ghaemmaghmi et al., 2016). RST that includes time-based delays progressively increase the proportional time spent without reinforcement regardless of behavior (Saini et al., 2016). To illustrate, suppose a child gains functional communication skills during initial skill-based intervention, a signaled period can be introduced, such as a red board to indicate that requests will not be honored for a period of time. During this time, the child can exhibit any other behavior, functional or otherwise, and once that programmed interval elapses, all communication will again be reinforced. As the child begins to tolerate brief delays to reinforcement, the programmed delay is slowly extended until the majority of time spent is in the context without reinforcement. While seemingly convenient for clinicians—time-based delays allow the clinician to set a fixed timer and deliver reinforcers irrespective of child behavior—several studies have found that challenging behavior is likely to reemerge during the delays when no behavioral expectations are imposed (e.g. Briggs et al., 2018; Muething et al., 2021).

Contingency-based delays attempt to account for the lack of behavioral expectations by targeting skill repertoires during the delay. Rather than passively awaiting a prespecified period of time before delivering reinforcement, contingency-based delays require the individual to engage in some expected behavior (i.e. skills) in order to produce that same reinforcement. The behavioral expectations of the child during the contingency-based delay can be programmed to vary in three possible ways (Iannaccone & Jessel, 2021). First, a negative contingency can be arranged whereby reinforcers will be returned following the absence of challenging behavior. In such scenarios, the child is often given the option to engage with other activities that are independently available while a caregiver is busy. The child need not necessarily engage with those activities because the behavioral expectation is simply to refrain from exhibiting challenging behavior. The duration of the delay is then progressively increased until the child

is able to wait appropriately for a caregiver for a specified amount of time. This contingency has been termed a differential reinforcement of other (DRO) behavior procedure and will hereby be described as a DRO-based delay (Weston et al., 2018).

Second, the contingency-based delay can be programmed to include a positive contingency that promotes a level of cooperation. That is, the reinforcers are only returned after a directed specified task is completed without challenging behavior. The number of tasks the child is required to complete is then progressively increased until the proportion of work meets the socially appropriate levels specified by the caregivers or educators (Hanley et al., 2014). This contingency has been termed a differential reinforcement of alternative (DRA) behavior procedure and will hereafter be described as a DRA-based delay. The behavioral expectations can then be combined in the third contingency-based delay variation that programs for reinforcement to be returned following both the absence of challenging behavior and cooperation with an adult (i.e. both DRO- and DRA-based delay; Jessel, Ingvarsson, Metras, Kirk & Whipple, 2018). Regardless of the arrangement, contingency-based delays are fundamentally different from time-based delays in that how long the child has to wait for the reinforcers to be returned is dependent on behavior during that time.

It is important to point out that, while autistic children are often the recipients of behavioral interventions including RST, this does not infer that autistic adolescents and adults will not benefit from similar services (Coffey et al., 2020; Ghaemmaghmi et al., 2016). For example, Ghaemmaghmi et al. (2016) conducted a within-subject comparison of time-based and contingency-based delays including four participants who exhibited challenging behavior. One of the participants was a nonverbal, 30-year-old male who had an extensive history of SIB and participated in a day habilitation program. During the contingency-based delay, the participant was prompted to complete a certain number of tasks, such as beading blocks on a string, before the reinforcers were returned. Including this positive contingency resulted in the greatest reductions in challenging behavior in comparison to the other time-based procedures he also experienced. Although the tasks to be completed during the delay were dependent on the participant's developmental abilities, extensions of RST to more complex vocational tasks seem well within the range of RST procedures to aid autistic individuals who may be entering the workforce.

To date, there have been multiple empirical demonstrations of the efficacy of RST using time-based delays (e.g. Muharib, Voggt, et al., 2021; Quigley et al., 2021) and contingency-based delays (e.g. Jessel, Ingvarsson, Metras, Whipple, et al., 2018; Rose & Beaulieu, 2019). In addition, meta-analyses have been conducted evaluating specific procedural variations of RST (Muharib et al., 2019; Muharib, Walker, et al., 2021); however, there has yet to be

a meta-analysis focusing on contingency-based delays. That may be due to the fact that many contingency-based procedures have only been recently developed within the past 10 years. Therefore, a meta-analysis is critically warranted given the rapid advancement of contingency-based procedures and may improve the efficacy of behavioral services provided to individuals with ASD who exhibit challenging behavior.

Meta-analyses offer a strategy to investigate the effectiveness of contingency-based delay procedures with the underlying aim of increasing the efficacy of applied behavior analytic services (see Dowdy, Peltier, et al., 2021 for a review). Furthermore, meta-analyses can help to identify effectiveness and expose areas where more research is needed on skill-based interventions that can be used by community-based providers and caregivers within naturalistic settings. The purpose of this review is to conduct a multilevel meta-analysis of RST procedures using contingency-based delays and summarize qualitative characteristics regarding the procedures while calculating obtained effect sizes. An omnibus effect size of contingency-based delays on challenging behavior is calculated along with relative effect sizes given different moderators (e.g. age, diagnosis, and procedural variation). The inclusion of moderators in the meta-analysis will help to define the conditions under which the contingency-based delays are predicted to be the most effective. The five research questions we attempted to answer were as follows:

- (1) What are the characteristics of studies that involved a contingency-based delay following functional communication training?
- (2) Is RST with contingency-based delays efficacious for autistic individuals compared with autistic individuals who have additional diagnoses or non-autistic individuals but with other diagnoses?
- (3) What is the differential effectiveness of contingency-based delays using positive, negative, or combined contingencies?
- (4) What is the differential effectiveness of DRA-based delays compared with DRO-based delays?
- (5) What are the notable moderators that could contribute to the efficacy of contingency-based delays?

## Method

### Search procedure

We systematically searched seven electronic databases to locate potential studies. These were Academic Search Ultimate, EBSCO, Education Resources Information Clearinghouse (ERIC), PsychINFO, Medline, Open Dissertations, and Psychology and Behavioral Sciences. We used a set of predetermined keywords divided into two lines within the online library searches. The first line included keywords related to FCT (i.e. functional communication

training OR communication training OR functional communication). The second line included keywords related to schedule thinning (i.e. delay to reinforcement OR delay OR reinforcer delay OR schedule thinning OR reinfor\* schedule thinning OR schedule of reinforcement OR contingency-based delay\* OR tolerance training OR chained schedules of reinforcement OR demand fading OR skill-based intervention). We used the term “AND” across the two lines to obtain all possible studies resulted from the combinations of terms. We restricted our searches to studies published in English. Next, we used the “Cited By” function on Google Scholar to identify potential studies that cited Hanley et al. (2014). Finally, we conducted ancestral searches by reviewing the reference lists of four relevant published literature reviews (Ghaemmaghami et al., 2021; Hagopian et al., 2011; Muharib et al., 2019; Neely et al., 2018). We completed all searches in November 2020 which resulted in a total of 3581 studies after removing duplicates (see Figure 1 for a flowchart).

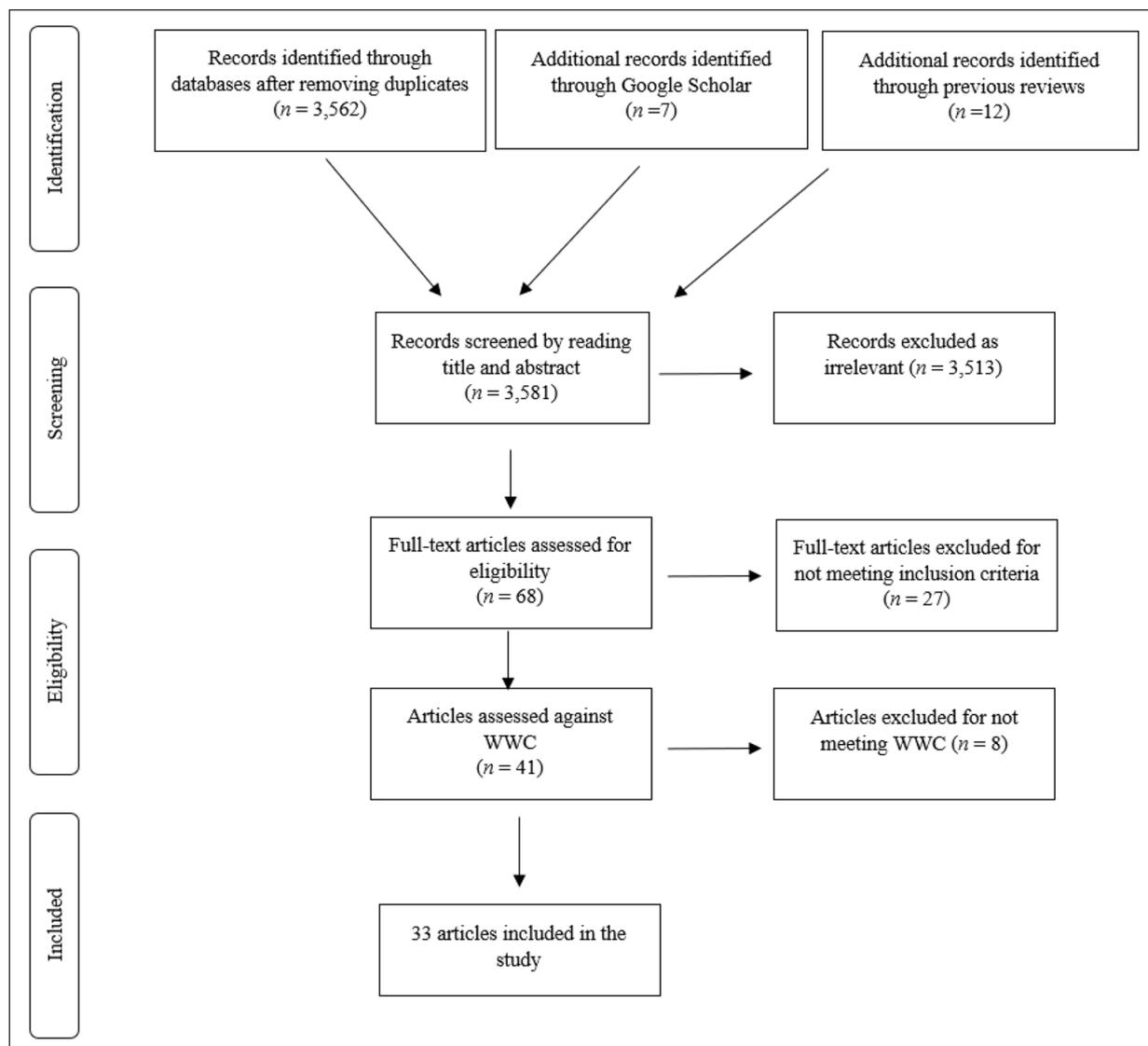
### Inclusion and exclusion criteria

Inclusion criteria were developed before the literature search. To be included in this meta-analysis, studies must have used an experimental research design that allowed for analysis of intervention effect on participant behavior (i.e. group design or single-case experimental design). We evaluated each study against the following inclusion criteria: (a) the intervention consisted of FCT and then a contingency-based delay to reinforcement which we defined as delays of delivering a reinforcer after an individual emits an FCR, and the delivery of the requested reinforcer is dependent on either a passage of time with no challenging behavior (i.e. DRO) and/or performance of specific tasks (i.e. DRA); (b) the study employed an experimental design; and (c) the study received a rating of a Meet or Meet with Reservations on the What Works Clearinghouse standards (WWC) (refer to Kratochwill et al., 2013). We excluded studies that did not meet one or more of the aforementioned criteria. We reviewed the abstracts of the 3581 potential studies. We then accessed the full text of 68 studies, which were not excluded based on an abstract review, to apply the first two inclusion criteria. This led to a total of 41 studies before further evaluations using the WWC standards.

Next, we evaluated each of the 41 studies against WWC standards. Based on this evaluation, we excluded 8 studies that did not meet the standards and included the remaining 33 studies (30 published studies, and 3 unpublished dissertations and theses) in this review.

### Data extraction and coding

We extracted descriptive information across each of the 67 participants represented in the 33 included studies in terms of (a) participant characteristics (age and diagnosis); (b)



**Figure 1.** PRISMA search process.

Note. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) summary of article extraction process.

dependent measures (targeted behaviors and functions of behavior); (c) type of contingency and response requirement; (d) duration of the initial delay and terminal delay; (e) number of tasks in the initial delay and terminal delay; and (f) intervention dosage. We coded the data using “1” to indicate the variable was relevant to the participant or “0” to indicate the variable was not relevant to the participant.

**Participant characteristics.** We included five categories for the age variable (0–5 years, 6–10 years, 11–15 years, 16–20 years, and 21 or older). For the diagnosis variable, we included three categories and coded each participant as having (a) autism only (i.e. autism, autistic disorder, pervasive developmental disorder not otherwise specified, pervasive developmental disorder, or Asperger’s); (b) autism + defined as being diagnosed with autism in addition to a

secondary diagnosis such as intellectual disability (ID), developmental or speech delay, attention deficit hyperactivity disorder (ADHD), oppositional defiant disorder (ODD), apraxia, fetal alcohol syndrome, or mood disorder; or (c) no autism defined as any diagnosis rather than autism such as ID, ADHD, or ODD.

**Dependent measures.** We coded the topography of the targeted challenging behavior for each participant. These included (a) aggression (such as hitting, kicking, and throwing objects at others); (b) self-injury (such as head-banging, skin picking, and pica); (c) property destruction (such as throwing objects not directed at others and breaking objects); (d) disruption (such as crying, screaming, protesting, spitting, and climbing on furniture); (e) elopement (i.e. leaving a designated area); and (f) disrobing as

described by the original authors as targeted for the participants. We also coded the function of the challenging behavior as reported by the authors of the original studies. These were (a) positive, (b) negative, or (c) both positive and negative.

**Type of contingency and response requirement.** We coded the type of contingency for each participant as (a) DRA-based wherein a reinforcer is given contingent on completing a number of predetermined skills/tasks; (b) DRO-based wherein a reinforcer is given contingent on not engaging in challenging behavior for a pre-determined period of time; or (c) both DRA and DRO-based contingencies in place at the same time. In addition, we coded the response requirement for each participant as (a) not engaging in challenging behavior only, (b) engaging in an activity/task, and (c) engaging in a tolerance response (e.g. saying “OK”).

**Duration of the initial and terminal delay.** For participants whose intervention was partly or fully time-based, we coded the initial and terminal delays. The initial delays were 0–5 s, 11–15 s, 16–30 s, or longer than 1 min. The terminal delays were 11–30 s, 1–2 min, 3–5 min, 6–12 min, or 28–34 min. These codes were created based on the information presented in the original studies. As an example, we did not include 31–60 s as a code for the terminal delay because none of the participants experienced such delay.

**Number of tasks in the initial delay and terminal delay.** For participants whose intervention was partly or fully DRA-based, we coded the number of tasks required to complete in the initial and terminal delays. The numbers in the initial delay were 1–2 tasks or 5–7 tasks. The numbers in the terminal delay were 1–4 tasks, 5–10 tasks, 15–25 tasks, or 30–42 tasks. Again, these codes were created based on the information presented in the original studies. As an example, we did not include 11–14 tasks as a code for the terminal delay because none of the participants experienced such response requirement.

**Intervention dosage.** We coded whether intervention was implemented 1–2 days per week or 3–5 days per week for each participant. For participants whose intervention dosage was unclear, we coded it as “cannot determine.” In addition to these coded variables, we also extracted additional data to be reported descriptively. These were related to the settings wherein the intervention was implemented (e.g. clinic, school) as well as the interventionist (e.g. researcher and parent).

### **Interrater reliability**

**Inclusion.** Since we narrowed it down to 68 full-text articles, we used this pool of articles to assess interrater reliability (IRR) for inclusion. To determine IRR for inclusion,

a second reviewer independently reviewed a randomly selected 30% portion of the 68 studies ( $n=20$ ) with reliability demonstrated at 100% agreement.

**WWC.** Because we had 41 studies that met our inclusion criteria before applying WWC criteria, we used this pool of articles to assess IRR for WWC evaluations. To determine the IRR for WWC evaluations, a second reviewer independently evaluated randomly selected 34% of the 41 studies ( $n=14$ ) against WWC standards. Our IRR was 92.8%. We discussed our disagreement related to Fisher et al. (2018) and resolved the disagreement which resulted in 100% agreement.

**Coding.** A second reviewer independently coded variables for 25 (37%) of the 67 participants to determine the IRR for coding the variables. Our agreement was 100%.

### **Data extraction and statistical analysis**

**Data extraction.** Webplotdigitizer, a web-based application plot digitizing tool available to users free of charge, was used to extract single-case experimental design (SCED) data for all included SCED studies (Rohatgi, 2015). Webplotdigitizer has been shown to be a reliable and valid tool for extracting data from SCED graphs (Drevon et al., 2017). In this meta-analysis, both baseline and intervention were extracted from the line graphs of each included study. A task analysis for using Webplotdigitizer was followed during data extraction by both the primary (Dowdy) and secondary coder (PhD student) who both had prior experience using Webplotdigitizer for extracting single-case experimental design data.

**Effect size estimation.** Given that all studies that met our inclusionary criteria were SCED studies, the log response ratio (LRR; Pustejovsky, 2018) effect size was calculated for each individual participant at the case level which has been shown to meet assumptions of SCED data (Pustejovsky, 2019). The LRR effect size is the natural log of proportionate change in the mean level between conditions and has been estimated in several recent meta-analyses that included SCED data (e.g. Bruhn et al., 2020; Dowdy et al., 2020). The LRR effect size has been shown to be appropriate for outcomes measured on a ratio scale such as rate, duration, and frequency measurement, and not appropriate for outcomes measured on a rating scale (Pustejovsky, 2018). Data included in this meta-analysis were drawn from single-case research design studies that were reported on a ratio scale. Autocorrelation is considered to be the degree of correlation between data points that are serially dependent on each other due the repeated nature of single-case research design and has been a challenge when synthesizing single-case research (Ledford & Gast, 2018). Barnard-Brak et al. (2021) found that the LRR effect size

has been shown to be well suited to address concerns of autocorrelation, thus influencing our decision to estimate the LRR effect size in this meta-analysis.

In its basic parametric form, the formula of LRR can be defined as

$$\varphi = \ln(\mu_{\beta}) - \ln(\mu_{\alpha})$$

where  $\mu_{\alpha}$  serves as the mean level outcome during the baseline phase,  $\mu_{\beta}$  serves as the mean level outcome during the intervention phase, and  $\ln$  denotes the natural logarithm function. If there were no change in level between phases, the LRR estimate is  $\varphi = 0$ .

Two versions of the LRR are available for estimating effect sizes, one version designed to synthesize studies in which behavioral outcomes are expected to decrease (e.g. reduction of challenging behavior) named LRRd, and the other version designed to synthesize studies in which behavior outcomes are expected to increase (e.g. teaching prosocial behavior). We selected to use LRRd because all of the dependent variables were considered to be challenging behaviors that were aimed at reducing. All LRRd estimates were calculated in the R (R Core Team, 2021) using the SingleCaseES package (Pustejovsky & Swan, 2018).

**Meta-analysis.** For the included studies, estimated LRRd effect sizes were then synthesized using a multilevel random effects model meta-analysis at the level of the participant and the level of the study. A multilevel approach was selected because we were interested in the extent to which the nesting nature and efficacy of contingency-based schedule thinning varied across individuals with ASD, studies, and selected moderators. This heterogeneity can be captured using a multilevel modeling meta-analytic approach (Becraft et al., 2020; Moeyaert et al., 2020). Effect size estimates were meta-analyzed using the metafor package in the R statistical environment (Viechtbauer, 2010). Standard deviations at the participant level characterized the degree of heterogeneity across participants at the study level; whereas standard deviations at the study level characterized the degree of heterogeneity across studies due to differences between settings, implementation, outcomes, or inclusion criteria. Restricted maximum likelihood (REML) was used to pool effect sizes in our model. Cluster-robust variance estimation (CRVE; Pustejovsky & Tipton, 2018) was used to estimate the overall effects across studies and participants. CRVE is a small sample bias correction designed to account for potential inaccuracy of effect size standard errors due to autocorrelation. The ClubSandwich (Pustejovsky, 2018) package in the R statistical environment was used to apply the CRVE correction.

**Publication bias.** In our effort to account for publication bias, unpublished dissertations and theses ( $n=3$ ; 9%) were included in this meta-analysis (Dowdy, Hantula, et al., 2021). If a dissertation or thesis was later published, the

published version of the study was included. To assess for publication bias we first visually inspected for the presence of asymmetry of LRRd effect size estimates at the study level in our funnel plot included in our supplemental files. We then statistically tested for publication bias by using the Egger's regression test (Egger et al., 1997). Visual analysis of the funnel plot used to detect asymmetry did not suggest the possibility of publication bias, and the Egger's test did not result in significant publication bias ( $p=0.23$ ). In our effort to promote open science (Hales et al., 2019; Hantula, 2019), raw data and the R scripts for replicating our analysis are available via Supplemental Materials (Muharib et al., 2021).

## Results

### Descriptive findings

Research question 1 involved summarizing the characteristics of studies that involved a contingency-based delay following functional communication training. Descriptive findings of participant characteristics are presented in Tables 1 and 2. In this analysis, 67 participants received a contingency-based delay following FCT. Over half of the participants (53.7%) had a diagnosis of ASD only and 32.9% of participants had a diagnosis of ASD and an additional diagnosis. The remaining 13.4% did not have a diagnosis of ASD. Over half of the participants fell in the age groups of 0–5 years and 6–10 years (29.9%, 35.8%, respectively). In terms of targeted challenging behaviors, aggression was the most common (71.6%) followed by property destruction (55.2%). Challenging behavior was maintained by positive reinforcement for 26.9% of the participants, negative reinforcement for 23.9% of the participants, and both positive and negative reinforcement for 46.3% participants.

For a majority of participants (54 out of 67), the contingency-based delay was DRA-based. For 82% of those participants, the response requirement involved a completion of a task or a number of tasks. In the initial delay, the number of tasks typically ranged between 1 and 2 tasks (58.2% of participants), whereas in the terminal delay, the most occurring range of number of tasks was 15–25 tasks (20% of participants). By contrast, for participants whose reinforcement delay was time-based, the most frequent duration of the initial delay was 0–5 s (22.3% of participants) and most frequent duration of the terminal delay was 3–5 min (18.1% of participants). A total of 39.3% of participants received intervention 3–5 times a week whereas 4.5% of participants received intervention less frequently (1–2 times per week). For 38 participants (56.2%), information about the intervention dosage was not specified.

### Meta-analysis findings

Effect size estimates were calculated to answer Research questions 2–5.

**Table 1.** Descriptive summaries of the included studies.

Article	Participants (name, age in years, diagnoses)	Targeted challenging behaviors	Target function(s)	Type of schedule thinning	Duration or number of task in terminal delay	Interventionist	Setting	WWC
Beaulieu et al. (2018)	Ally, 7, ASD	Aggression, SIB, disruption	Positive; tangible	DRA-based	28–34 min	BCBA	Home	Meet w/rev
Bennett (2018)	Mario, 5, ASD Luigi, 6, ASD	Disruption, elopement	Both positive and negative; escape and tangibles (Mario) Positive; tangibles (Luigi)	DRA-based	30–43 task (Mario); 6–12 min (Luigi)	Therapist	Home	Meet w/rev
Boesch et al. (2015)	Mike, 14, ASD	SIB	Positive; tangible	DRA-based	1–2 min	Researcher	Classroom	Meet w/rev
Boyle et al. (2020)	Aaron, 8, ASD	Elopement	Positive; Access to stereotypy	DRO-based	6–12 min	Therapist	University building	Meet
Briggs et al. (2018)	Samantha, 12, ASD, ID	Aggression, SIB, property destruction	Both positive and negative; tangible, escape	DRA-based	15–25 tasks	Researcher	Clinic	Meet
Carr & Carlson (1993)	Mark, 18, ASD, ID Bob, 17, ASD, ID	Aggression, SIB, disruption	Both positive and negative; escape, tangible (Mark and Bob) Positive; tangible (Danny)	DRA-based	1–4 tasks	Staff	Community (grocery store)	Meet
Carr et al. (1999)	Val, 14, ASD, ADHD Gary, 17, ASD, ID Juan, 38, ASD, ID	Aggression, disruption (Val) Aggression, SIB, disruption (Gary and Juan)	Negative; escape (Gary) not specified (Val and Juan)	DRA-based	1–2 larger tasks	Caregivers	Home	Meet w/rev
Coffey et al. (2020)	Ryan, 9, ASD Alan, 10, ASD, FAS, ADHD, mood disorder	Aggression, property destruction, disruption (Ryan, Alan), disrobing (Ryan), elopement (Alan)	Both positive and negative; escape, tangible/activity	DRA-based	5–10 tasks	Staff (generalization probes by parents and teachers)	Clinic, home (generalization)	Meet w/rev
Drifke et al. (2020)	Logan, 11, ASD, ADHD Andy, 14, ASD, ADHD, schizencephaly Tanner, 9, ID	Aggression, property destruction	Positive; tangible (Logan, Tanner), attention (Andy)	DRA-based, DRO-based	3–5 min (Logan and Andy); 1–2 min (Tanner); 30–42 tasks (Andy)	Therapist	Home	Meet w/rev
Falconata, Roane, et al. (2012)	Steph, 8, ASD Mike, 8, ASD	Aggression, property destruction, disruption, disrobing (Steph); Aggression, SIB (Mike)	Both positive and negative; escape, preferred activity	DRA-based	6–12 min	Therapist	Self-contained room	Meet w/rev
Falconata, White, & Muehling (2012)	Danny, 8, ASD	Disruption, property destruction	Both positive and negative; attention, escape, tangible	DRA-based	1–4 tasks	Therapist	Self-contained classroom	Meet w/rev
Falconata et al. (2013)	Allonzo, 7, ASD Joe, 12, ASD	Aggression, disruption (Allonzo); Aggressive, SIB (Joe)	Both positive and negative; escape, tangible	DRA-based	3–5 min	Therapist	Self-contained room	Meet w/rev

(Continued)

**Table 1.** (Continued)

Article	Participants (name, age in years, diagnoses)	Targeted challenging behaviors	Target function(s)	Type of schedule thinning	Duration or number of task in terminal delay	Interventionist	Setting	WWC
Ferguson et al. (2020)	Heather, 8, ASD	Aggression, disruption, property destruction	Positive; tangibles, mand compliance	DRA-based	15–25 tasks	Researcher	Clinic, home (parent training)	Meet w/rev
Fisher et al. (2018)	Erica, 16, ASD, ADHD Corey, 3, ASD Jaden, 8, ASD Derek, 7, ASD	Aggression, SIB (all); Property destruction (Corey, Jaden, Derek)	Both positive and negative; escape and tangible (Erica and Derek); Positive; tangibles, attention (Corey); tangible (Jaden) Negative; escape	DRO-based	1–30 s (Erica, Cory, Derek); 1–2 min (Jaden)	Therapist, Caregiver (Corey)	Clinic	Meet
Gerow et al. (2020)	Noah, 6, ASD, ADHD Liam, 4, ASD, speech delay	Aggression	Negative; escape	DRA-based	5–10 tasks (Noah); 15–25 task (Liam)	Therapist (doctoral student)	Clinic	Meet
Ghaemmaghami et al. (2016)	Will, 30, ASD, ID, ADHD, mood disorder, Jack, 21 months, ASD Alex, 6, ASD	Aggression, property destruction, elopement (Jack, Alex); SIB (Will)	Positive; access to edibles (Will), Positive; attention, tangible (Jack). Both positive and negative; attention, escape, mand compliance (Alex)	DRA-based & DRO-based (Will); combination of DRA and DRO-based (Jack and Alex)	3–5 min (Jack and Will); 6–12 min (Alex); 5–10 tasks Will	Researcher	Clinic (Jack, and Alex); rehabilitation center (Will)	Meet w/rev
Greer et al. (2019)	Kari, 3, ASD	Aggression, SIB, property destruction	Both positive and negative; attention, escape	DRA-based	5–10 tasks	Parent, caregiver	Classroom, clinic	Meet w/rev
Hagopian et al. (1998)	Case 19, ID	Aggression, SIB, disruption	Both positive and negative; attention, escape	DRA-based	5–10 tasks	Therapist	Inpatient unit	Meet w/rev
Hanley et al. (2014)	Gail, 3, ASD Dale, 11, ASD Bob, 8, ASD	Aggression, disruption, property destruction	Positive; attention, tangible (Gail); mand compliance (Dale) Both positive and negative; tangible/activity, escape (Bob)	DRA-based	Unspecified	Researcher (all), Parent (Gail)	Clinic (all), classroom (Dale)	Meet w/rev
Jessel, Ingvarsson, Metras, Kirk, & Whipple (2018)	Joe, 5, ASD Kane, 5, ASD, ADHD, ID John, 4, ASD Steve, 4, ASD Zane, 10, ASD, ADHD	Aggression, disruption Elopement	Negative; escape Positive; tangibles/activity	DRA-based (John); DRO-based (Joe); DRA and DRO-based (Kane)	3–5 min (Joe, Kane); 5–12 tasks (Kane); 30–42 tasks (John)	Therapist	Clinic	Meet w/rev
Jessel, Ingvarsson, Metras, Whipple, et al. (2018)	Bill, 28, ASD, ID Maggie, 30, ASD, ID Michael, 26, ASD, ID	Aggression, SIB (all), disruption (Bill, Michael), property destruction (Bill); SIB (Joe, Jen), Aggression (Kim)	Negative; escape (Bill, Michael) Both positive and negative; tangible, escape (Maggie) Negative; escape	DRA-based	1–4 tasks	Staff	Community	Meet
Lalli et al. (1995)	Joe, 10, ID Jen, 15, ID, ASD Kim, 13, ID, ASD	SIB (Joe, Jen), Aggression (Kim)	Negative; escape	DRA-based	15–25 tasks	Therapist	Hospital	Meet
Mildon et al. (2004)	R, 4, ASD	Aggression, disruption, property destruction	Negative; escape	DRA-based	5–10 tasks	Researcher	Home	Meet w/rev

(Continued)

Table 1. (Continued)

Article	Participants (name, age in years, diagnoses)	Targeted challenging behaviors	Target function(s)	Type of schedule thinning	Duration or number of task in terminal delay	Interventionist	Setting	WWC
Nuhu (2016)	Nancy, 4, ASD Adam, 6, apraxia and deletion of chromosome 5.0, Derek, 5, DD Kevin, 5, ASD Teddy, 9, ID	Aggression (Kevin, Derek), disruption (Adam, Nancy)	Negative; escape (Nancy, Adam, Derek) Both positive and negative; escape, tangible (Kevin)	DRA-based	15–25 tasks	Therapist	Clinic (therapy room in university campus)	Meet w/rev
Peck Peterson et al. (2005)	Teddy, 9, ID	Property destruction	Both positive and negative; escape, tangible	DRA-based	15–25 tasks	Teachers (undergrad and grad students)	Classroom	Meet w/rev
Perry & Fisher (2001)	Anna, 12, ID, ODD	Aggression, SIB, property destruction	Both positive and negative; escape, tangible	DRA-based	1–4 tasks		Hospital	Meet
Rose & Beaulieu (2019)	Anna, 3, ASD Owen, 5, ASD	Disruption (Anna); Aggression, disruption, property destruction (Owen)	Positive; tangible, attention	DRA and DRO-based	6–12 min	Therapist	Home	Meet w/rev
Santiago et al. (2016)	Zeke, 14, ASD Karen, 11, ASD	Aggression, SIB, disruption (Zeke), Aggression (Karen)	Both positive and negative; escape, tangible, attention (Zeke) Positive; tangible, attention (Karen)	DRA-based	1–2 min	Researcher; teacher (treatment extension for Zeke) Parent	Classroom (Zeke), Home (Karen)	Meet w/rev
Suess (2015)	Scott, 6, ASD Cade, 6, ASD Mark, 4, ASD Derick, 3, ASD	Aggression, SIB, disruption, property destruction	Both positive and negative; escape, tangible (Cade, Scott, Mark) Negative; escape (Derick)	DRA-based	1–4 tasks		Home	Meet w/rev
Taylor et al. (2018)	Adam, 12, ASD, ADHD, dyspraxia	Aggression, disruption, property destruction	Both positive and negative; escape, tangible	DRA-based	28–34 min	BCBA, teachers (treatment extension)	School	Meet w/rev
Zangrillo et al. (2016)	Cody, 7, ASD, IED Matt, 7, ADHD, IED	Aggression, property destruction (both), disruption (Matt)	Negative; escape	DRA-based (Matt), Combination of DRA- and DRO-based (Cody)	3–5 min	Therapist	Clinic	Meet

WWC: What Works Clearinghouse; ASD: autism spectrum disorder; SIB: self-injurious behavior; ADHD: attention deficit hyperactivity disorder; DRA: alternative-based delay; DRO: differential reinforcement of other; DD: defiant disorder; ODD: oppositional defiant disorder; BCBA: Board Certified Behavior Analyst; FAS: Fetal Alcohol Syndrome; IED: Intermittent Explosive Disorder.

**Table 2.** Primary characteristics of the 67 participants.

Category	Subcategory	% of participants
Diagnosis	ASD diagnosis only	53.7
	ASD diagnosis +	32.9
	No ASD	13.4
Age	0–5 years	29.9
	6–10 years	35.8
	11–15 years	19.4
	16–20 years	7.45
	21 years and older	7.45
Targeted challenging behavior	Aggression	71.6
	Property destruction	55.2
	Disruption	49.2
	Self-injury	40.2
	Elopement	8.9
	Disrobing	3.0
Function of behavior	Positive	26.9
	Negative	23.9
	Both positive and negative	46.3
	Not specified	2.9
Type of contingency	DRA-based	80.5
	DRO-based	11.9
	Both	11.9
Response requirement	Engaging in a task	82.0
	Engaging in a vocal tolerance response (e.g. “OK”)	23.8
	No challenging behavior only	20.8
Duration of the initial delay	0–5 s	22.3
	11–15 s	3.0
	16–30 s	7.6
	Longer than 1 min	4.4
Duration of the terminal delay	11–30 s	4.4
	1–2 min	7.4
	3–5 min	18.1
	6–12 min	10.6
	28–34 min	3.0
	30–42 min	6.1
Number of tasks in the initial delay	1–2 tasks	58.2
	5–7 tasks	1.4
Number of tasks in the terminal delay	1–4 tasks	17.9
	5–10 tasks	10.4
	15–25 tasks	20.0
	30–42 tasks	6.1
Treatment dosage	1–2 days per week	4.5
	3–5 days per week	39.3
	Not specified	56.2

ASD: autism spectrum disorder, DRA: differential reinforcement of alternative behavior, DRO: differential reinforcement of other behavior.

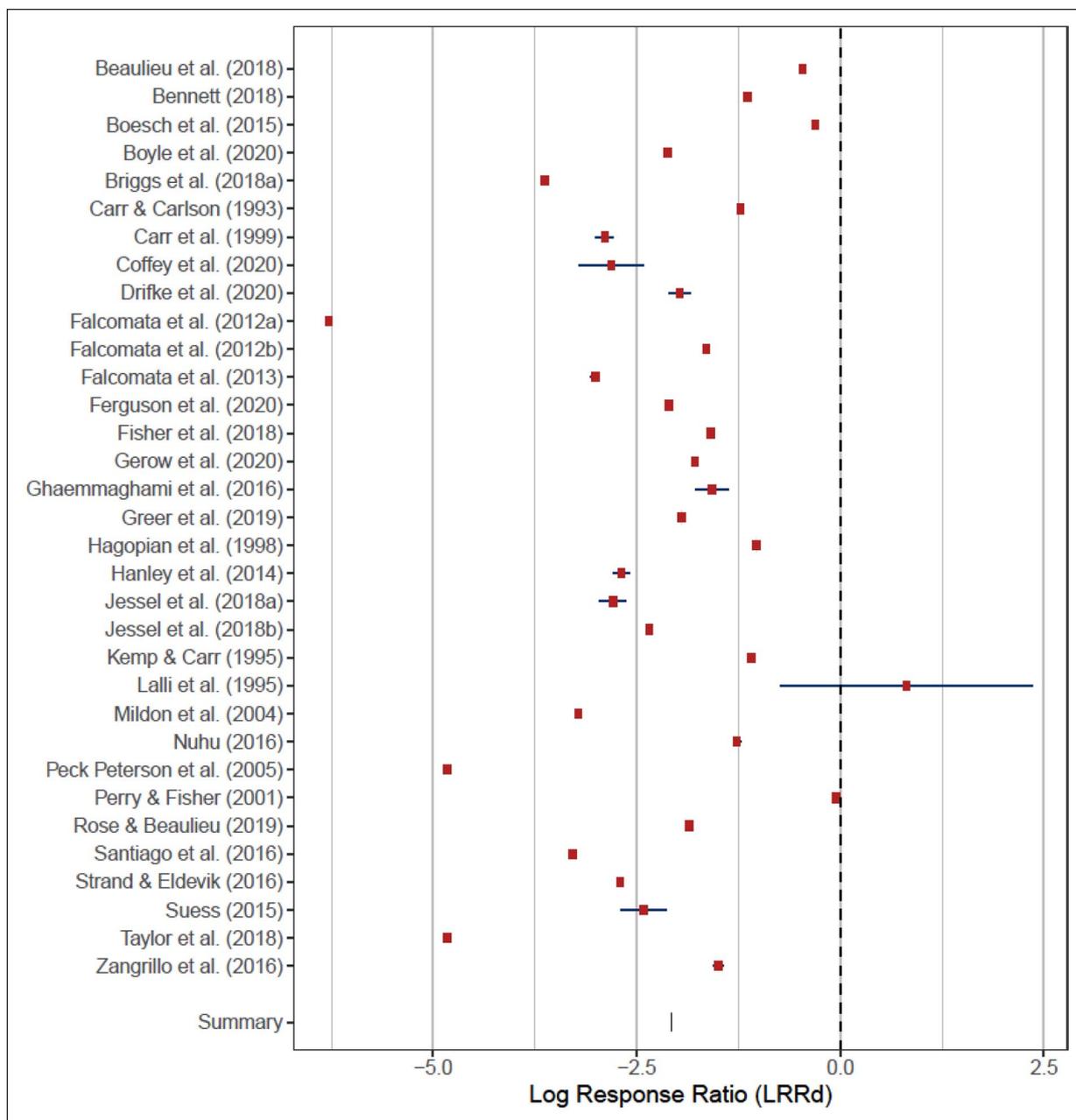
Numbers do not add up to 100% in type of contingency because three participants were exposed to DRA and DRO separately, so for them we coded, DRA-based and DRO-based.

Numbers do not add up to 100% in response requirement as some participants were required to engage in a vocal tolerance response and engage in an alternative activity, so they were coded in both variables.

Numbers do not add up to 100% duration of initial/terminal delay, number of tasks in the initial/terminal delay, and treatment dosage due to a lack of specific information for some participants in the original studies.

Heterogeneity was estimated using  $I^2$  which resulted in 94% of the total variance for the model. In addition, 24% of the total variance estimated using  $I^2$  was due to between-cluster heterogeneity, whereas 69% of total variance

estimated was due to within-cluster heterogeneity. Figure 2 shows a forest plot of LRRd effect size estimates and 95% confidence intervals per study included in the meta-analysis. Effect size estimates, standard errors, degrees of freedom,



**Figure 2.** Forest plot of LRRd effect size estimates and 95% confidence intervals.

and confidence intervals for moderating variables are presented in Table 3.

Research question 2 examined whether RST using contingency-based delays is efficacious for autistic individuals as compared with individuals with autism plus other diagnoses and individuals without an autism diagnosis. Our findings show that this approach appears to be most efficacious for those with an ASD diagnosis ( $LRRd = -2.17$ ; 95%  $CI = (-2.76, -1.58)$ ) when compared with ASD plus other diagnoses ( $LRRd = -2.14$ ; 95%  $CI = (-2.80, -1.50)$ ) and no diagnosis ( $LRRd = -1.95$ ; 95%  $CI = (-3.55, -0.35)$ ). However, all subcategories of diagnosis show favorable

outcomes to suggest that RST using contingency-based delays is an efficacious approach to treat challenging behavior for individuals who do and do not have an ASD diagnosis.

Research question 3 examined the differential effectiveness of contingency-based delays using positive, negative, or combined contingencies. Our findings show that contingencies appear to have the greatest effect when they incorporate positive reinforcement. Findings showed that positive reinforcement-based contingencies had the greatest effect size ( $LRRd = -2.30$ ; 95%  $CI = (-2.83, -1.78)$ ) followed by contingencies that included both positive and

**Table 3.** Log response ratio decreasing (LRRd) effect size per moderating variable.

Category	Effect size estimates of moderating variables			
	LRRd ES	SE	df	95% CI
<b>Diagnosis</b>				
ASD diagnosis only	-2.17	0.28	21.41	[-2.76, -1.58]
ASD diagnosis+	-2.14	0.31	16.04	[-2.80, -1.50]
No ASD	-1.95	0.66	6.49	[-3.55, -0.35]
<b>Age</b>				
0–5 years	-2.32	0.21	16.11	[-2.77, -1.87]
6–10 years	-2.29	0.33	21.40	[-2.97, -1.61]
11–15 years	-1.71	0.59	9.24	[-3.03, -0.38]
16–20 years	-1.66	0.48	2.17	[-3.55, 0.24]
21 years and older	-2.19	0.65	2.19	[-4.77, 0.40]
<b>Targeted challenging behavior</b>				
Aggression	-1.21	0.52	12.36	[-2.33, -0.08]
Self-injurious behavior	0.01	0.56	12.90	[-1.20, 1.21]
Disruption	-0.23	0.36	17.21	[-0.98, 0.53]
Elopement	-1.84	0.51	2.93	[-3.47, -0.21]
Property destruction	-1.02	0.49	14.54	[-2.07, 0.24]
Disrobing	-0.60	1.51	1.35	[-11.26, 10.08]
<b>Behavior contingencies</b>				
Positive	-2.30	0.24	11.66	[-2.83, -1.78]
Negative	-1.74	0.36	7.27	[-2.59, -0.89]
Both positive and negative	-2.21	0.31	21.90	[-2.86, -1.55]
<b>Type of contingency</b>				
DRO-based	-1.24	0.60	1.97	[-3.84, 1.37]
DRA-based	-2.13	0.29	27.14	[-2.72, -1.55]
Both	-1.74	0.52	3.87	[-3.19, -0.29]
<b>Alternative response during thinning</b>				
No challenging behavior only	-0.42	0.34	4.81	[-1.31, 0.46]
Engaging in an alternative task	-1.81	0.29	22.53	[-2.41, -1.22]
Engaging in a vocal tolerance response (e.g. "OK")	-0.98	0.45	8.49	[-2.01, 0.60]
<b>Duration of the terminal delay</b>				
1–2 min	-1.34	0.89	2.31	[-4.72, 2.04]
3–5 min	-0.89	0.81	4.53	[-3.05, 1.27]
6–12 min	-1.05	0.77	2.52	[-3.80, 1.70]
28–34 min	-2.63	2.19	1.00	[-30.40, 25.13]
<b>Number of tasks in the terminal delay</b>				
1–4 tasks	-1.30	0.38	3.99	[-2.37, -0.24]
5–10 tasks	-1.51	0.38	3.68	[-2.61, -0.41]
15–25 tasks	-0.53	0.70	5.68	[-2.26, 1.21]
30–42 tasks	-1.22	1.10	2.32	[-5.39, 2.94]
<b>Dosage of treatment</b>				
1–2 days per week	-0.82	0.34	1.00	[-5.11, 3.48]
3–5 days per week	-1.87	0.32	12.88	[-2.57, -1.18]

ASD: autism spectrum disorder; DRA: differential reinforcement of alternative behavior; DRO: differential reinforcement of other behavior; SE: standard error; CI: confidence interval.

Cluster-robust variance estimation to correct for small sample size and autocorrelation was applied to each moderating variable and as presented in the table.

negative reinforcement (LRRd=-2.21; 95% CI=(-2.86, -1.55)). The negative reinforcement based subcategory resulted in the smallest effect size (LRRd=-1.74; 95% CI=(-2.59, -0.89)), but still presented favorably (i.e. reduced challenging behavior), thus suggesting that RST

using contingency-based delays appear effective irrespective of reinforcement type.

Research question 4 examined the differential effectiveness between DRO-based delays compared with DRA-based delays. As reported in Table 3, our findings showed

that DRA-based delays were overall more effective and presented with a smaller standard error (LRRd = -2.13; 95% CI = (-2.72, -1.55)) when compared with DRO-based delays (LRRd = -1.24; 95% CI = (-3.84, 1.37)). However, all delay types (DRO-based, DRA-based, and both) that were included in this meta-analysis resulted in favorable behavior change which highlight their effectiveness.

Research question 5 aimed to identify notable moderators that may contribute to the efficacy of contingency-based delays. Based upon statistics reported in Table 3, noteworthy findings show that RST using contingency-based delays are effective irrespective of age and appear to be most effective when targeting elopement (LRRd = -1.84; 95% CI = (-3.47, -0.21)) and aggression (LRRd = -1.21; 95% CI = (-2.33, -0.08)) as a topography of challenging behavior. Furthermore, our findings show that the most effective intervention dosage is three to five days per week (LRRd = -1.87; 95% CI = (-2.57, -1.18)) compared with 1–2 days per week (LRRd = -0.82; 95% CI = (-5.11, 3.48)).

## Discussion

We sought to summarize and meta-analyze studies addressing challenging behavior that evaluated RST procedures with contingency-based delays. In studies that employed experimental designs commensurate with WWC standards (Kratochwill et al., 2013), we found that contingency-based RST procedures following FCT were evaluated across 67 participants, over half of whom had an autism diagnosis. Contingency-based RST procedures have been applied more commonly with children than adolescents or adults but have been demonstrated to be efficacious in treating multiple different topographies of challenging behavior, sensitive to a variety of reinforcement contingencies, and across varying magnitudes of delay to reinforcement. Favorable outcomes are likely to be achieved with contingency-based RST procedures, with marginally better outcomes observed with individuals diagnosed only with autism, notably better outcomes observed when positive reinforcement was identified in the contingency maintaining challenging behavior (i.e. positive and combined contingencies), and notably better outcomes obtained with DRA-based as opposed to DRO-based RST. Finally, moderator analyses suggested that contingency-based RST procedures appear to be more effective when implemented three-to-five times per week as opposed to less frequently.

### *Characteristics of studies that involved contingency-based RST*

Overall, RST procedures involving a contingency-based delay were evaluated across a broad range of participant

profiles and intervention contexts. However, the majority of evaluations (66%) involved children as opposed to adolescents or adults. Autistic adults and adults with related disabilities often require a similar intensity of services, and are just as much in need of evidence-based interventions with outcomes sustainable under practical schedules of reinforcement and in natural contexts. Furthermore, as autistic adults enter the workforce, it is important that they are able to access services that teach toleration and cooperation when presented with job-related expectations in the absence of important reinforcers. Therefore, it is important for future researchers to continue to evaluate RST procedures with adolescent and adult populations to better understand the conditions under which they may prove most beneficial.

One encouraging takeaway from the descriptive summary in the current study is that it appears that critical stakeholders (e.g. parents, teachers) implemented RST procedures in ecologically valid contexts with contingency-based delays more commonly than has been reported in previous reviews of RST (e.g. Muharib et al., 2019). The factors that may have contributed to this increase are currently unclear. Perhaps researchers have taken to heart recent calls to emphasize external and ecological validity in behavioral research with autistic individuals, and have thus made a concerted effort to expand upon the known efficacy of RST procedures (e.g. Marchand et al., 2011; Smith et al., 2007). Ghaemmghami et al. (2021) argued that extension of positive treatment outcomes to ecologically relevant contexts was a critical step to establish the effectiveness of reinforcement-based procedures such as FCT, noting that a large majority of published demonstrations involved implementation by highly-trained experimenters. We argue that similar evidence is critical for RST procedures as well, and encourage more research involving application of FCT and RST procedures by relevant caregivers. Increased application of procedures implemented and validated by parents and teachers will not only help clarify the degree of external validity of such procedures, but it may help future researchers refine procedures so that they are tailored to the needs of the individual client and the relevant caregivers in their life. For example, Galpin et al. (2018) found that many caregivers desire a family-centered approach to service provision to help caregivers address “specific needs and characteristics of their autistic child” (p. 576). It can be difficult and stressful for parents to cope with challenging behavior when their child cannot functionally communicate (Schiltz et al., 2018), and this is possibly exacerbated when service providers implement services without properly supporting the extension of those services to relevant caregivers.

The current review found that relevant caregivers implemented intervention in 8 of the 32 (25%) reviewed studies. By contrast, Muharib et al. (2019) found that

caregivers implemented RST procedures in only 4 of the 29 (14%) studies they reviewed. Intervention implementation by a relevant caregiver often represents the natural environment that RST procedures are designed to emulate; future research should examine the relative efficacy and social acceptability of various RST procedures when exclusively implemented by teachers and parents to progress toward practical, feasible intervention of challenging behavior. It is worth noting that a limitation of the current review, relevant to the concerns reported by Galpin et al. (2018), is that we did not code or incorporate information regarding the social acceptability or validity of RST procedures and the outcomes they produce. Regardless of differential efficacy, or the observation that more caregivers were trained to implement RST with contingency-based delays relative to that which has been reported in other reviews of RST, conclusions regarding which type of procedure is more favorable are incomplete in the absence of information pertaining to social validity.

### *Efficacy of contingency-based RST for autistic individuals relative to other diagnoses*

RST with contingency-based delays has proven efficacious for individuals with ASD only, with ASD and additional diagnoses, and without ASD but with a different clinical diagnosis. Applications for autistic individuals collectively showed marginally greater effect sizes, suggesting particular utility for this population. One reason for this may be because of the specific skill deficits associated with ASD, such as communication deficits (Weiss et al., 2019), difficulties with inhibitory control (Christ et al., 2007), and functional engagement in contextually appropriate behavior (Lanovaz et al., 2013). Although the overall findings suggest that one's diagnostic profile does not appear to be a precluding factor to effective intervention, the differential utility of contingency-based RST procedures for autistic individuals is encouraging, considering that challenging behavior when exhibited by this population constitutes one of largest barriers to improved QoL (Ruef & Turnbull, 2002; van Heijst & Geurts, 2015) and one of the most notable contributors to parental mental health issues (Weiss et al., 2012).

### *Differential efficacy of contingency-based delays using positive, negative, or combined contingencies*

RST with contingency-based delays were efficacious in reducing challenging behavior sensitive to all types of contingencies. These results show that contingencies involving positive reinforcement (either alone or in combination with negative reinforcement) produced notably larger overall effect sizes than contingencies involving only negative reinforcement. One could interpret an RST procedure

involving DRA-based delay as necessarily involving negative reinforcement during the period in which an individual experiences positive reinforcement. This is because DRA-based RST requires that the individual engage in contextually appropriate behavior during a delay to reinforcement that is usually incompatible with consumption behavior. For example, if a child was asked to complete an academic task following functional communication in order to earn time with their preferred activities (i.e. DRA-based RST with positive reinforcement), implicit in the teaching arrangement is the notion that the positive reinforcers are not delivered until the task requirement is met, at which point the instruction is terminated (i.e. probable negative reinforcement). When the individual is engaging with positive reinforcers, it is unlikely that alternative expectations are in place. By contrast, RST involving only negative reinforcement suggests that the individual is provided a break from expectations without alternative activities to occupy time. Taken together, these findings suggest that clinicians should consider incorporating positive reinforcement into RST procedures for negatively reinforced behavior to achieve greater reductions in challenging behavior. For example, during RST procedures in Zangrillo et al. (2016), positive reinforcers were not only incorporated for individuals with challenging behavior found to be maintained by negative reinforcement, but they were necessary in order to achieve long-term reductions in challenging behavior.

### *Comparison of DRA- and DRO-based RST*

In addition to confirming the overall efficacy of contingency-based RST procedures in reducing challenging behavior, we found differentially greater effect sizes from evaluations of DRA-based RST relative to DRO-based RST. This means that RST procedures that specify and require alternative behavior during periods of nonreinforcement may yield stronger intervention outcomes than procedures that merely require that challenging behavior not occur during nonreinforcement periods. This finding is supported by several studies that have shown that teaching alternative skills to be used in moments where challenging behavior may have been likely has the capacity to create durable, generalizable improvements in challenging behavior (Durand & Carr, 1991, 1992; Hernandez et al., 2007). In other words, teaching a skill repertoire to replace challenging behavior has the potential to reduce challenging behavior in contexts beyond the one in which the skills were taught.

One reason that DRA-based RST may have yielded stronger treatment outcomes is that teaching an individual to engage in specific alternative responses during delays to reinforcement enables them to produce reinforcers with their own behavior, which we interpret as providing the individual with a greater level of control and agency during potentially evocative periods (Hanley et al., 1997;

Luczynski & Hanley, 2009). For example, if a learner's challenging behavior is reinforced by preferred foods, teaching the learner to request food items allows them to obtain the food when they are hungry as opposed to following some arbitrary time period without food. Furthermore, if their initial request to obtain food is met with a denial or delay cue (e.g. "I can't make that for you right now"), teaching them specific cooperative responses to engage in may help mediate the delay by not only putting the learner in a position to control when food is delivered (e.g. "I can make that for you after you make your bed and wash your hands"), but by increasing the opportunity for alternative behavior to contact other reinforcers. This finding has further implications for the capacity of RST-procedures to help improve self-advocacy skills in individuals with ASD; teaching alternative responses empowers individuals to advocate for their wants and needs without resorting to challenging behavior.

We also note that the terminal delay, whether DRA-based or DRO-based or both, varied across studies. It is difficult to associate magnitude of delay with the efficacy of the procedure because there is no standardized approach to determining terminal intervention goals. Many factors may have influenced the terminal intervention schedule, including the nature of the research question, the length of one's participation in a study, the individuals involved in deciding the terminal goal, and that which is actually expected of the individual in their natural environment. Therefore, it is premature to conclude which type of contingency-based RST procedure is likely to yield greater periods of nonreinforcement without challenging behavior. Future research should directly compare the duration of delay that is achievable with either type of contingency-based RST procedure; although, we suspect that DRA-based procedures would be capable of chaining larger delays due to expected engagement in alternative activities. Furthermore, although various research questions might lend themselves to particular terminal schedules in either DRA- or DRO-based RST procedures, it is important to consider the naturalistic conditions that such procedures are designed to ultimately emulate (this was not overtly taken into consideration in many of the reviewed studies). It may be the case that relevant stakeholders, not researchers, are best suited to decide how "thin" a terminal schedule should be based upon their everyday circumstances. Future RST research should more explicitly and preemptively articulate (a) what the terminal treatment goal is and (b) who contributed to determining this goal. Such information will elucidate the extent to which certain RST procedures are capable of achieving ecologically relevant schedules of reinforcement, and a collection of such findings may influence the trajectory of still more research examining the effectiveness of RST procedures in addressing challenging behavior in naturalistic environments.

### *Moderators that contribute to the efficacy of contingency-based delays*

Moderator analyses were conducted to investigate if, and to what extent, variability across studies could be explained by conceptually relevant variables. In a recent review of moderators in meta-analyses of single-case research design, Moeyaert et al. (2021) shared that current best practice is to use a multilevel modeling approach to account for heterogeneity. Despite this recommendation, they found that few meta-analyses of single-case research design in recent years had adopted a multilevel approach when modeling how variables moderated intervention effectiveness. Here, we adopted Moeyaert et al. (2021)'s recommendation of modeling potential moderating variables using a multilevel approach; furthermore, metaregression models consisting of cluster-robust variance estimation and small sample corrections were applied to account for small samples (Tipton & Pustejovsky, 2015). Although we included modeling techniques designed for small samples and selected to use a multilevel meta-analytic approach, an increased sample size resulting from future research on contingency-based delays to reinforcement would strengthen the precision of moderator outcomes (Jamshidi et al., 2020).

RST with contingency-based delays appears to have generality of effects across age groups, topographies of challenging behavior, and intervention dosages. Among the topographies of challenging behavior that we coded for shown in Table 3, RST with contingency-based delays are most efficacious when targeting elopement and aggression. By contrast, RST with contingency-based delays yielded less of a magnitude of effect when SIB was targeted. It is important to note that multiple studies involved intervention of a group of topographies of challenging behavior simultaneously, with rates typically reported as aggregated measures including SIB. Because target behavior topographies were seldom separated in graphs presented in the reviewed studies, conclusions regarding topography as a moderator of the efficacy of RST are difficult to make. However, future research can more closely examine the factors that may improve the efficacy of RST interventions for socially mediated SIB.

RST procedures showed larger effects when interventions were delivered 3–5 days per week compared with 1–2 days per week. This finding suggests that more therapeutic visits per week tend to result in better outcomes. Our findings align with previous research on early intensive behavioral interventions for autistic children, which emphasize contingency-based skill building in a manner similar to RST procedures with contingency-based delays, and have also been demonstrated to be more effective in the acquisition of some skills with higher dosage (Eldevik et al., 2009; Granpeesheh et al., 2009). Future research

may want to consider coding for different dimensions of intervention dosage (e.g. number of sessions and number of trials) to more precisely pinpoint what exactly is resulting in the improvement in challenging behavior during intensive behavioral intervention.

## Conclusion and implications for practice

Autistic individuals are indeed at greater risk of exhibiting dangerous challenging behavior (Adamek et al., 2011), which is likely to become a chronic concern that negatively impacts QoL in the absence of intervention (Chiang & Wineman, 2014). Although previous literature reviews have highlighted the efficacy of behavioral interventions like FCT (Ghaemmaghami et al., 2021) and certain procedural variations of RST (Hagopian et al., 2011; Muharib et al., 2019; Muharib, Walker, et al., 2021), none have focused on the effects of contingency-based delays in sustaining positive outcomes for autistic individuals and related disabilities.

Recent single-subject comparative analyses have revealed more favorable outcomes—with respect to challenging behavior reduction and replacement skill acquisition—when RST involved contingency-based delays as opposed to time-based delays (Drifke et al., 2020; Ghaemmaghami et al., 2016). Furthermore, Iannaccone and Jessel (2021) found that contextually appropriate behavior was more likely to be acquired and maintained when contingency-based RST procedures were DRA-based as opposed to DRO-based. Therefore, given the somewhat recent emergence of novel contingency-based delay procedures, the finding that studies evaluating intensive behavioral interventions for individuals with ASD tend to report only short-term effects (Rodgers et al., 2021), and the understanding that RST is critical to the sustained success of behavioral interventions supporting the needs of individuals with ASD (Galpin et al., 2018; Ghaemmaghami et al., 2021; Hagopian et al., 2011), the current meta-analytic review contributes directly to the literature to illustrate that contingency-based RST procedures have produced robust and durable intervention outcomes across heterogeneous intervention contexts. Specifically, behavioral interventions that attempt to address challenging behavior in a manner that can accommodate an individual's natural environment are more successful when they (a) are implemented often, (b) include specific behavioral expectations of the individual during periods of nonreinforcement, and (c) involve positive reinforcement.

## Declaration of conflicting interests

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

## Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

## Informed consent

For this type of study formal consent is not required.

## Ethical approval

This article does not contain any studies with human participants performed by any of the authors.

## ORCID iDs

Art Dowdy  <https://orcid.org/0000-0001-6466-6774>

Joshua Jessel  <https://orcid.org/0000-0002-1649-2834>

## Note

1. “Functional” refers to any events that are determined to be relevant based upon a functional assessment.

## References

\* Denotes that the research study was included in the meta-analysis

- Adamek, L., Nichols, S., Tetenbaum, S. P., Bregman, J., Ponzio, C. A., & Carr, E. G. (2011). Individual temperament and problem behavior in children with autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities*, 26(3), 173–183. <https://doi.org/10.1177/1088357611405041>
- Muharib, R., Dowdy, A., Rajaraman, A., & Jessel, J. (2021). Contingency-based delay to reinforcement following functional communication training for autistic individuals: A multilevel meta-analysis [Supplemental Files]. [https://osf.io/p5736/?view\\_only=e52ab176b4804463986cb8f9afebc3eb](https://osf.io/p5736/?view_only=e52ab176b4804463986cb8f9afebc3eb)
- Baker, J. K., Fenning, R. M., Howland, M. A., & Huynh, D. (2019). Parental criticism and behavior problems in children with autism spectrum disorder. *Autism*, 23(5), 1249–1261. <https://doi.org/10.1177/1362361318804190>
- Barnard-Brak, L., Watkins, L. & Richman, D.M. (2021). Autocorrelation and estimates of treatment effect size for single-case experimental design data. *Behavioral Interventions*, 36(3), 595–605. <https://doi.org/10.1002/bin.1783>
- \*Beaulieu, L., Van Nostrand, M. E., Williams, A. L., & Herscovitch, B. (2018). Incorporating interview-informed functional analyses into practice. *Behavior Analysis in Practice*, 11(4), 385–389. <https://doi.org/10.1007/s40617-018-0247-7>
- Becraft, J. L., Borrero, J. C., Sun, S., & McKenzie, A. A. (2020). A primer for using multilevel models to meta-analyze single case design data with AB phases. *Journal of Applied Behavior Analysis*, 53(3), 1799–1821. <https://doi.org/10.1002/jaba.698>
- \*Bennett, B. (2018). *Chained schedule thinning to increase compliance within a functional communication training context* [Unpublished master's thesis, St. Cloud State University].
- \*Boesch, M. C., Taber-Doughty, T., Wendt, O., & Smalts, S. S. (2015). Using a behavioral approach to decrease self-injurious behavior in an adolescent with severe autism: A data-based

- case study. *Education and Treatment of Children*, 38(3), 305–328. <https://doi.org/10.1353/etc.2015.0012>
- \*Boyle, M. A., Bacon, M. T., Brewer, E. M., Carton, S. M., & Gaskill, L. A. (2020). Evaluating a treatment without extinction for elopement maintained by access to stereotypy. *Journal of Applied Behavior Analysis*, 53(3), 1531–1541. <https://doi.org/10.1002/jaba.682>
- \*Briggs, A. M., Akers, J. S., Greer, B. D., Fisher, W. W., & Retzlaff, B. J. (2018). Systematic changes in preference for schedule-thinning arrangements as a function of relative reinforcement density. *Behavior Modification*, 42(4), 472–497. <https://doi.org/10.1177/0145445517742883>
- Bruhn, A., Gilmour, A., Rila, A., Van Camp, A., Sheaffer, A., Hancock, E., Fernando, J., & Wehby, J. (2020). Treatment components and participant characteristics associated with outcomes in self-monitoring interventions. *Journal of Positive Behavior Interventions*. Advance online publication. <https://doi.org/1098300720946651>
- \*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., Levin, L., McConnachie, G., Carlson, J. I., Kemp, D. C., Smith, C. E., & McLaughlin, D. M. (1999). Comprehensive multi situational 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>
- Chiang, H. M., & Wineman, I. (2014). Factors associated with quality of life in individuals with autism spectrum disorders: A review of literature. *Research in Autism Spectrum Disorders*, 8(8), 974–986. <https://doi.org/10.1016/j.rasd.2014.05.003>
- Christ, S. E., Holt, D. D., White, D. A., & Green, L. (2007). Inhibitory control in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 37(6), 1155–1165. <https://doi.org/10.1007/s10803-006-0259-y>
- \*Coffey, A. L., Shawler, L. A., & Jessel, J. (2020). Interview-informed synthesized contingency analysis (IISCA): Novel interpretations and future directions. *Behavior Analysis in Practice*, 13(1), 217–225. <https://doi.org/10.1007/s40617-019-00348-3>
- Dowdy, A., Hantula, D. A., Travers, J. C., & Tincani, M. (2021). Meta-analytic methods to detect publication bias in behavior science research. *Perspectives on Behavior Science*. Advance online publication. <https://doi.org/10.1007/s40614-021-00303-0>
- Dowdy, A., Peltier, C., Tincani, M., Schneider, W. J., Hantula, D. A., & Travers, J. C. (2021). Meta-analyses and effect sizes in applied behavior analysis: A review and discussion. *Journal of Applied Behavior Analysis*, 54(4), 1317–1340. <https://doi.org/10.1002/jaba.862>
- Dowdy, A., Tincani, M., & Schneider, W. J. (2020). Evaluation of publication bias in response interruption and redirection: A meta-analysis. *Journal of Applied Behavior Analysis*, 53(4), 2151–2171. <https://doi.org/10.1002/jaba.724>
- Drevon, D., Fursa, S. R., & Malcolm, A. L. (2017). Intercoder reliability and validity of WebPlotDigitizer in extracting graphed data. *Behavior Modification*, 41(2), 323–339. <https://doi.org/10.1177/0145445516673998>
- \*Drifke, M. A., Tiger, J. H., & Lillie, M. A. (2020). DRA contingencies promote improved tolerance to delayed reinforcement during FCT compared to DRO and fixed-time schedules. *Journal of Applied Behavior Analysis*, 53(3), 1579–1592. <https://doi.org/10.1002/jaba.704>
- 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>
- Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ (Clinical research ed.)*, 315(7109), 629–634. <https://doi.org/10.1136/bmj.315.7109.629>
- Eldevik, S., Hastings, R. P., Hughes, J. C., Jahr, E., Eikeseth, S., & Cross, S. (2009). Meta-analysis of early intensive behavioral intervention for children with autism. *Journal of Clinical Child & Adolescent Psychology*, 38(3), 439–450. <https://doi.org/10.1080/15374410902851739>
- \*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>
- \*Falcomata, T. S., Roane, H. S., Muething, C. S., Stephenson, K. M., & Ing, A. D. (2012). Functional communication training and chained schedules of reinforcement to treat challenging behavior maintained by terminations of activity interruptions. *Behavior Modification*, 36(5), 630–649. <https://doi.org/10.1177/0145445511433821>
- \*Falcomata, T. S., White, P., & Muething, C. S. (2012). A functional communication training and chained schedule procedure to treat challenging behavior with multiple functions. *Journal of Developmental and Physical Disabilities*, 24(6), 529–538. <https://doi.org/10.1007/s10882-012-9287-z>
- \*Ferguson, J. L., Leaf, J. A., Cihon, J. H., Milne, C. H., Leaf, J. B., McEachin, J., & Leaf, R. (2020). Practical functional assessment: A case study replication and extension with a child diagnosed with autism spectrum disorder. *Education and Treatment of Children*, 43, 171–185. <https://doi.org/10.1007/s43494-020-00015>
- \*Fisher, W. W., Greer, B. D., Fuhrman, A. M., Saini, V., & Simmons, C. A. (2018). Minimizing resurgence of destructive behavior using behavioral momentum theory. *Journal of Applied Behavior Analysis*, 51(4), 831–853. <https://doi.org/10.1002/jaba.499>
- Galpin, J., Barratt, P., Ashcroft, E., Greathead, S., Kenny, L., & Pellicano, E. (2018). ‘The dots just don’t join up’: Understanding the support needs of families of children on the autism spectrum. *Autism*, 22(5), 571–584. <https://doi.org/10.1177/1362361316687989>

- \*Gerow, S., Radhakrishnan, S., Davis, T. N., Hodges, A., & Feind, A. (2019). A comparison of demand fading and a dense schedule of reinforcement during functional communication training. *Behavior Analysis in Practice, 13*(1), 90–103. <https://doi.org/10.1007/s40617-019-00403-z>
- Ghaemmaghani, 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>
- Ghaemmaghani, M., Hanley, G. P., & Jessel, J. (2021). Functional communication training: From efficacy to effectiveness. *Journal of Applied Behavior Analysis, 54*(1), 122–143. <https://doi.org/10.1002/jaba.762>
- Ghaemmaghani, M., Hanley, G. P., Jessel, J., & Landa, R. (2018). Shaping complex functional communication responses. *Journal of Applied Behavior Analysis, 51*(3), 502–520. <https://doi.org/10.1002/jaba.468>
- Granpeesheh, D., Dixon, D. R., Tarbox, J., Kaplan, A. M., & Wilke, A. E. (2009). The effects of age and treatment intensity on behavioral intervention outcomes for children with autism spectrum disorders. *Research in Autism Spectrum Disorders, 3*(4), 1014–1022. <https://doi.org/10.1016/j.rasd.2009.06.007>
- 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*(1), 4–16. <https://doi.org/10.1007/BF03391770>
- \*Hagopian, L. P., Fisher, W. W., Sullivan, M. T., Acquisto, J., & LeBlanc, L. A. (1998). Effectiveness of functional communication training with and without extinction and punishment: A summary of 21 inpatient cases. *Journal of Applied Behavior Analysis, 31*(2), 211–235. <https://doi.org/10.1901/jaba.1998.31-211>
- Hales, A. H., Wesselmann, E. D., & Hilgard, J. (2019). Improving psychological science through transparency and openness: An overview. *Perspectives on Behavior Science, 42*(1), 13–31. <https://doi.org/10.1007/s40614-018-00186-8>
- \*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>
- Hantula, D. A. (2019). Editorial: Replication and reliability in behavior science and behavior analysis: A call for a conversation. *Perspectives on Behavior Science, 42*(1), 1–11. <https://doi.org/10.1007/s40614-019-00194-2>
- Hernandez, E., Hanley, G. P., Ingvarsson, E. T., & Tiger, J. H. (2007). A preliminary evaluation of the emergence of novel mand forms. *Journal of Applied Behavior Analysis, 40*(1), 137–156. <https://doi.org/10.1901/jaba.2007.96-05>
- Hume, K., Steinbrenner, J. R., Odom, S. L., Morin, K. L., Nowell, S. W., Tomaszewski, B., Szendrey, S., McIntyre, N. S., Yücesoy-Özkan, S., & Savage, M. N. (2021). Evidence-based practices for children, youth, and young adults with autism: Third generation review. *Journal of Autism and Developmental Disorders, 51*(11), 4013–4032. <https://doi.org/10.1007/s10803-020-04844-2>
- Iannaccone, J. A., & Jessel, J. (2021). A translational comparison of contingency-based progressive delay procedures and their effects on contextually appropriate behavior. *Journal of Applied Behavior Analysis, 54*(1), 231–247. <https://doi.org/10.1002/jaba.780>
- Jamshidi, L., Declercq, L., Fernández-Castilla, B., Ferron, J. M., Moeyaert, M., Beretvas, S. N., & Van den Noortgate, W. (2020). Multilevel meta-analysis of multiple regression coefficients from single-case experimental studies. *Behavior Research Methods, 52*(5), 2008–2019. <https://doi.org/10.3758/s13428-020-01380-w>
- \*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>
- \*Jessel, J., Ingvarsson, E. T., Metras, R., Whipple, R., Kirk, H., & Solsbery, L. (2018). Treatment of elopement following a latency-based interview-informed, synthesized contingency analysis. *Behavioral Interventions, 33*(3), 271–283. <https://doi.org/10.1002/bin.1525>
- 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*(7), 926–937. <https://doi.org/10.1007/s10803-010-1118-4>
- \*Kemp, D. C., & Carr, E. G. (1995). Reduction of severe problem behavior in community employment using a hypothesis-driven multicomponent intervention approach. *Journal of the Association for Persons with Severe Handicaps, 20*(4), 229–247. <https://doi.org/10.1177/154079699602000401>
- Kratochwill, T. R., Hitchcock, J. H., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2013). Single-case intervention research design standards. *Remedial and Special Education, 34*(1), 26–38. <https://doi.org/10.1177/0741932512452794>
- \*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>
- Lanovaz, M. J., Robertson, K. M., Soerono, K., & Watkins, N. (2013). Effects of reducing stereotypy on other behaviors: A systematic review. *Research in Autism Spectrum Disorders, 7*(10), 1234–1243. <https://doi.org/10.1016/j.rasd.2013.07.009>
- Ledford, J. R., & Gast, D. L. (Eds.) (2018). *Single case research methodology: Applications in special education and behavioral sciences* (3rd ed.). Routledge. <https://doi.org/10.4324/9781315150666>
- Leyfer, O. T., Folstein, S. E., Bacalman, S., Davis, N. O., Dinh, E., Morgan, J., Tager-Flusberg, H., & Lainhart, J. E. (2006). Comorbid psychiatric disorders in children with autism: Interview development and rates of disorders. *Journal of Autism and Developmental Disorders, 36*(7), 849–861. <https://doi.org/10.1007/s10803-006-0123-0>
- Lowe, K., Allen, D., Jones, E., Brophy, S., Moore, K., & James, W. (2007). Challenging behaviours: Prevalence

- and topographies. *Journal of Intellectual Disability Research*, 51(Pt. 8), 625–636. <https://doi.org/10.1111/j.1365-2788.2006.00948.x>
- Luczynski, K. C., & Hanley, G. P. (2009). Do children prefer contingencies? An evaluation of the efficacy of and preference for contingent versus noncontingent social reinforcement during play. *Journal of Applied Behavior Analysis*, 42(3), 511–525. <https://doi.org/10.1901/jaba.2009.42-511>
- Mandell, D. S. (2008). Psychiatric hospitalization among children with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 38(6), 1059–1065. <https://doi.org/10.1007/s10803-007-0481-2>
- 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>
- Matson, J. L., & Nebel-Schwalm, M. S. (2007). Comorbid psychopathology with autism spectrum disorder in children: An overview. *Research in Developmental Disabilities*, 28(4), 341–352. <https://doi.org/10.1016/j.ridd.2005.12.004>
- \*Mildon, R. L., Moore, D. W., & Dixon, R. S. (2004). Combining noncontingent escape and functional communication training as a treatment for negatively reinforced disruptive behavior. *Journal of Positive Behavior Interventions*, 6(2), 92–102. <https://doi.org/10.1177/10983007040060020401>
- Moeyaert, M., Manolov, R., & Rodabaugh, E. (2020). Meta-analysis of single-case research via multilevel models: Fundamental concepts and methodological considerations. *Behavior Modification*, 44(2), 265–295. <https://doi.org/10.1177/0145445518806867>
- Moeyaert, M., Yang, P., Xu, X., & Kim, E. (2021). Characteristics of moderators in meta-analyses of single-case experimental design studies. *Behavior Modification*. Advance online publication. <https://doi.org/10.1177/01454455211002111>
- Muething, C., Pavlov, A., Call, N., Ringdahl, J., & Gillespie, S. (2021). Prevalence of resurgence during thinning of multiple schedules of reinforcement following functional communication training. *Journal of Applied Behavior Analysis*, 54(2), 813–823. <https://doi.org/10.1002/jaba.791>
- Muharib, R., Alresheed, F., Ninci, J., Walker, V., & Voggt, A. P. (2019). Thinning schedules of reinforcement following functional communication training for children with intellectual and developmental disabilities: A meta-analytic review. *Journal of Autism and Developmental Disorders*, 49(12), 4788–4806. <https://doi.org/10.1007/s10803-019-04191-x>
- Muharib, R., Voggt, A. P., Wood, C. L., Pennington, R. C., & Lang, R. (2021). Further evaluation of a multiple schedule thinning procedure following functional communication training in a school setting. *Advances in Neurodevelopmental Disorders*, 5(2), 204–217. <https://doi.org/10.1007/s41252-021-00196-9>
- Muharib, R., Walker, V. L., Alresheed, F., & Gerow, S. (2021). Effects of multiple schedules of reinforcement on appropriate communication and challenging behaviors: A meta-analysis. *Journal of Autism and Developmental Disorders*, 51(2), 613–631. <https://doi.org/10.1007/s10803-020-04569-2>
- Neely, L., Garcia, E., Bankston, B., & Green, A. (2018). Generalization and maintenance of functional communication training for individuals with developmental disabilities: A systematic and quality review. *Research in Developmental Disabilities*, 79, 116–129. <https://doi.org/10.1016/j.ridd.2018.02.002>
- \*Nuhu, N. N. (2016). *Schedule thinning following functional communication training: A comparison of chained schedules and multiple schedules* [Unpublished master's thesis, Auburn University].
- \*Peck Peterson, S. M., Caniglia, C., Jo Royster, A., Macfarlane, E., Plowman, K., Jo Baird, S., & Wu, N. (2005). Blending functional communication training and choice making to improve task engagement and decrease problem behaviour. *Educational Psychology*, 25(2–3), 257–274. <https://doi.org/10.1080/0144341042000301193>
- \*Perry, A. C., & Fisher, W. W. (2001). Behavioral economic influences on treatment designed to decrease destructive behavior. *Journal of Applied Behavior Analysis*, 34(2), 211–215. <https://doi.org/10.1901/jaba.2001.34-211>
- Pustejovsky, J. E. (2018). Using response ratios for meta-analyzing single-case designs with behavioral outcomes. *Journal of School Psychology*, 68, 99–112. <https://doi.org/10.1016/j.jsp.2018.02.003>
- Pustejovsky, J. E. (2019). Procedural sensitivities of effect sizes for single-case designs with directly observed behavioral outcome measures. *Psychological Methods*, 24(2), 217–235.
- Pustejovsky, J. E., & Swan, D. M. (2018). *SingleCaseES: A calculator for single-case effect size indices* [R package version 0.4.0.9999]. <https://github.com/jepusto/SingleCaseES>
- Pustejovsky, J. E., & Tipton, E. (2018). Small-sample methods for cluster-robust variance estimation and hypothesis testing in fixed effects models. *Journal of Business & Economic Statistics*, 36(4), 672–683.
- Quigley, J., Dowdy, A., Trucksess, K., & Finlay, A. (2021). An investigation of functional communication training and schedule thinning using a multiple schedule on elopement to access stereotypy. *Journal of Autism and Developmental Disorders*, 51(9), 3224–3234. <https://doi.org/10.1007/s10803-020-04788-7>
- R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <http://www.R-project.org/>
- 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(4), 512–520. <https://doi.org/10.1007/s10803-011-1218-9>
- Richards, C., Oliver, C., Nelson, L., & Moss, J. (2012). Self-injurious behaviour in individuals with autism spectrum disorder and intellectual disability. *Journal of Intellectual Disability Research*, 56(5), 476–489. <https://doi.org/10.1111/j.1365-2788.2012.01537>
- Rodgers, M., Simmonds, M., Marshall, D., Hodgson, R., Stewart, L. A., Rai, D., . . . Coureur, A. L. (2021). Intensive behavioural interventions based on applied behaviour analysis for young children with autism: An international collaborative individual participant data meta-analysis. *Autism*, 25(4), 1137–1153. <https://doi.org/10.1177/1362361320985680>
- Rohatgi, A. (2015). *WebPlotDigitizer* (Version 3.9) [Computer software]. WebPlotDigitizer Online Software. <http://arohatgi.info/WebPlotDigitizer>

- \*Rose, J. C., & Beaulieu, L. (2019). 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>
- Ruef, M. B., & Turnbull, A. P. (2002). The perspectives of individuals with cognitive disabilities and/or autism on their lives and their problem behavior. *Research and Practice for Persons with Severe Disabilities, 27*(2), 125–140.
- Ruppel, K. W., Hanley, G. P., Landa, R. K., & Rajaraman, A. (2021). An evaluation of “balance”: A home-based, parent-implemented program addressing emerging problem behavior. *Behavior Analysis in Practice, 14*(2), 324–341. <https://doi.org/10.1007/s40617-020-00490-3>
- Saini, V., Greer, B. D., Fisher, W. W., Lichtblau, K. R., DeSouza, A. A., & Mitteer, D. R. (2016). Individual and combined effects of noncontingent reinforcement and response blocking on automatically reinforced problem behavior. *Journal of Applied Behavior Analysis, 49*(3), 693–698. <https://doi.org/10.1002/jaba.306>
- \*Santiago, J. L., Hanley, G. P., Moore, K., & Jin, C. S. (2016). The generality of interview-informed functional analyses: Systematic replications in school and home. *Journal of Autism and Developmental Disorders, 46*(3), 797–811. <https://doi.org/10.1007/s10803-015-2617-0>
- Schiltz, H. K., McVey, A. J., Magnus, B., Dolan, B. K., Willar, K. S., Pleiss, S., . . . Van Hecke, A. V. (2018). Examining the links between challenging behaviors in youth with ASD and parental stress, mental health, and involvement: Applying an adaptation of the family stress model to families of youth with ASD. *Journal of Autism and Developmental Disorders, 48*(4), 1169–1180. <https://doi.org/10.1007/s10803-017-3446-0>
- Simonoff, E., Pickles, A., Charman, T., Chandler, S., Loucas, T., & Baird, G. (2008). Psychiatric disorders in children with autism spectrum disorders: Prevalence, comorbidity, and associated factors in a population-derived sample. *Journal of the American Academy of Child and Adolescent Psychiatry, 47*(8), 921–929. <https://doi.org/10.1097/CHI.0b013e318179964f>
- 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>
- \*Suess, A. N. (2015). *An evaluation of generalization, renewal, and resurgence during functional communication training conducted via telehealth* [Doctoral dissertation, University of Iowa]. <https://doi.org/10.17077/etd.ysq8sulmy>
- \*Taylor, S. A., Phillips, K. J., & Gertzog, M. G. (2018). Use of synthesized analysis and informed treatment to promote school reintegration. *Behavioral Interventions, 33*(4), 364–379. <https://doi.org/10.1002/bin.1640>
- Tincani, M., Lorah, E., & Dowdy, . (2018). *Functional behavior assessment*. Pro-Ed.
- Tipton, E., & Pustejovsky, J. E. (2015). Small-sample adjustments for tests of moderators and model fit using robust variance estimation in meta-regression. *Journal of Educational and Behavioral Statistics, 40*(6), 604–634. <https://doi.org/10.3102/1076998615606099>
- van Heijst, B. F. C., & Geurts, H. M. (2015). Quality of life in autism across the lifespan: A meta-analysis. *Autism, 19*(2), 158–167. <https://10.1177/1362361313517053>
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*(3), 1–48. <https://doi.org/10.18637/jss.v036.i03>
- Weiss, J. A., Cappadocia, M. C., MacMullin, J. A., Vecili, M., & Lunsy, Y. (2012). The impact of child problem behaviors of children with ASD on parent mental health: The mediating role of acceptance and empowerment. *Autism: The International Journal of Research and Practice, 16*(3), 261–274. <https://doi.org/10.1177/13623613111422708>
- Weiss, M. J., Volpe, S., & Cihon, J. (2019). Autism: History and diagnosis. In Hupp S. & Jewell J. D. (Eds.), *The encyclopedia of child and adolescent development* (pp. 1–13). John Wiley & Sons, Inc. <https://doi.org/10.1002/9781119171492.wecad085>
- Weston, R., Hodges, A., & Davis, T. N. (2018). Differential reinforcement of other behaviors to treat challenging behaviors among children with autism: A systematic and quality review. *Behavior Modification, 42*(4), 584–609. <https://doi.org/10.1177/0145445517743487>
- \*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 Developmental Disabilities, 62*(3), 147–156. <https://doi.org/10.1080/20473869.2016.1176308>