Journal of Behavioral Education
https://doi.org/10.1007/s10864-021-09453-2

ORIGINAL PAPER

Reducing Escape without Escape Extinction: A Systematic Review and Meta-Analysis of Escape-Based Interventions

Kate T. Chazin[1] · Marina S. Velez[1] · Jennifer R. Ledford[1]

Accepted: 29 August 2021

Abstract
Individuals with disabilities may engage in challenging behavior to escape aversive stimuli, like academic tasks or non-preferred foods. Interventions to reduce these behaviors often employ escape extinction; that is, the implementer withholds escape following challenging behavior. Escape extinction can increase risk of injury, restrict autonomy, and worsen the learner–implementer relationship. To mitigate collateral effects, interventions can use strategies without escape extinction (i.e., escape-based); that is, implementers can provide escape contingent on challenging behavior during intervention, in conjunction with other intervention components. However, no comprehensive syntheses of these interventions have been conducted. We identified 39 articles that included escape-based interventions, which contained 273 single-case designs. Escape-based interventions were associated with lower levels of challenging behavior and higher levels of adaptive behavior than baseline conditions. Most comparisons between escape-based and escape extinction interventions showed no functional relation, indicating that escape extinction may not add substantial benefit to intervention efficacy.

Keywords Challenging behavior · Escape extinction · Extinction alternatives · Log response ratio · Single-case research design

Introduction

Individuals with disabilities are more likely to engage in challenging behavior to escape or avoid aversive stimuli (e.g., task directions, non-preferred foods) than for any other reason (Beavers et al., 2013; Zarcone et al., 1994a). Escape-maintained challenging behavior operates antithesis to a behavioral cusp. Opposite to the way

✉ Kate T. Chazin
kate.chazin@vanderbilt.edu

[1] Department of Special Education, Vanderbilt University, Peabody Box 228, Nashville, TN 37203, USA

Published online: 25 September 2021

that behavioral cusps open contact with new contingencies, reinforcers, and environments (Bosch & Fuqua, 2001), escape-maintained challenging behavior inhibits access to learning opportunities and produces cascading losses. For example, escape-maintained challenging behavior results in missed academic instruction (McComas et al., 2000), fewer social opportunities (Schmidt et al., 2013) and barriers to learning skills that promote health and independence (Schmidt et al., 2013; Wilder et al., 2005). These missed learning opportunities may explain poor adulthood outcomes correlated with challenging behavior in childhood, including academic dropout (Park & Scott, 2009), isolation from peers (Campbell et al., 2006), and low employment and income (Park & Scott, 2009). Conversely, interventions that reduce rates of challenging behavior improve long-term outcomes for individuals (Stormont et al., 2005); therefore, it is critically important to intervene to reduce these behaviors.

Escape extinction is one common strategy to reduce escape-maintained challenging behavior (Lane et al., 2007). In applying escape extinction, an implementer withholds escape contingent on challenging behavior (Cooper et al., 2020). This includes physical prompting, repeating task directions (paced prompting), and blocking escape attempts (Cooper et al., 2020). In a recent literature review on school-based treatments for escape-maintained challenging behavior, escape extinction was cited as the most common intervention component (Dart et al., 2018). Research indicates that escape extinction is an effective strategy for reducing escape-maintained challenging behavior (Geiger et al., 2010).

Despite its unquestionable efficacy, escape extinction is associated with collateral effects and social issues that may lead practitioners to consider alternative practices to escape extinction as the sole or primary intervention component (Geiger et al., 2010). We have identified five primary concerns with the use of escape extinction.

First, common topographies of escape extinction require provision of verbal, physical, or proximal attention. For learners with multiply-maintained challenging behavior (specifically, attention and escape), the implementer may reinforce challenging behavior intended for extinction (Borrero et al., 2010; Kern et al., 2002; Kodak et al., 2003). Even when attention is not a hypothesized reinforcer for challenging behavior, it may still serve as a competing stimulus that reduces efficacy of intervention (Gardner et al., 2009).

Second, authentic implementers (e.g., teachers, parents) are often the intended end users of interventions designed by behavior analysts, and they frequently implement interventions with imperfect treatment fidelity (Geiger et al., 2010; McConnachie & Carr, 1997). This is especially problematic with escape extinction, as intermittent reinforcement of challenging behavior (one possible lapse in treatment fidelity) can produce extinction-resistant challenging behavior (Cooper et al., 2020). This can decrease the efficacy of the intervention and worsen long-term outcomes.

Third, escape extinction procedures require that the implementer present aversive stimuli. When aversive stimuli are repeatedly paired with the instructional setting, materials, and implementer, these contextual stimuli can become conditioned punishers (Catania, 2013). When this occurs, challenging behavior is likely to persist in the presence of these contextual stimuli, even when the original aversive stimuli are not present (Iwata, 1987). Avoidance of instructional contexts directly impacts access to learning opportunities. Further, this conditioning process works in opposition to recommended procedures for developing rapport (Lugo et al., 2017).

Fourth, the use of restrictive procedures (i.e., those that limit an individual's freedom to move and act independently (UK Department of Health, 2015); e.g., full physical prompting) increases risk of harm to the learner and implementer, particularly when the learner struggles against physical prompts. Further, the use of restrictive procedures may not be feasible, acceptable, or sanctioned in some settings (e.g., schools).

Finally, restrictive procedures have been rated with low social acceptability among practitioners, caregivers, psychologists, and children receiving behavioral interventions (Elliot, 1988; Luiselli et al., 2015). Few data exist directly assessing the social acceptability of escape extinction specifically. However, in a recent direct assessment of intervention preferences, three children strongly preferred interventions incorporating differential reinforcement over those incorporating extinction (Owen et al., 2021), indicating that extinction may be non-preferred by direct consumers of behavior analytic interventions. Social acceptability is a core component of social validity (Wolf, 1978) and should be a consideration when selecting interventions.

Despite these side effects and social issues, escape extinction is recommended by behavioral experts (Geiger et al., 2010) and persists as a common component of interventions addressing escape-maintained challenging behavior in the research literature (Dart et al., 2018). This indicates a gap between current practice and practice that may be most ethical and socially valid. In the current review, we aim to address this gap by synthesizing the available research for interventions intended to reduce escape-maintained challenging behavior that do not include escape extinction. Hereafter, for brevity and clarity, we will refer to these interventions as escape-based interventions, in that escape is provided contingent on challenging behavior within intervention. However, it should be noted that all interventions included in this review include antecedent and consequent strategies beyond escape provision, and these strategies (not provision of escape contingent on challenging behavior) are the components theorized to result in treatment efficacy.

This study examined the efficacy of escape-based interventions, that is, interventions in which escape was provided contingent on challenging behavior. Despite the focus on escape-based interventions, this review benefits practitioners in that it provides strategies to *either* exclude escape extinction from treatment entirely (i.e., alternatives to escape extinction) *or* to add intervention components to escape extinction treatments to mitigate its side effects (i.e., complements to escape extinction). That is, practitioners can add antecedent modifications and changes to consequences following alternative behavior, whether or not escape extinction is an intervention component. In either case, the end goal is the same: to decrease restrictiveness and increase social validity of interventions intended to reduce escape-maintained challenging behavior. The research questions guiding the review were: (a) What escape-based interventions have been evaluated for individuals with escape-maintained challenging behavior? (b) What is the quality and rigor of this evidence base? (c) What were outcomes for visual and statistical analyses for identified escape-based interventions, compared to baseline conditions and escape extinction interventions?

Methods

Search Procedures

Two doctoral students enrolled in a special education program (i.e., the first and second authors; hereafter "we") completed all searching and coding. We searched for articles that included (a) participants who engaged in escape-maintained challenging behavior, as identified by a functional analysis (FA), (b) interventions designed to decrease challenging behavior, in which escape was provided contingent on challenging behavior (i.e., escape-based interventions), (c) comparisons between escape-based interventions and baseline conditions *or* between escape-based interventions and escape extinction interventions, (d) outcomes that included direct observational measures of challenging and/or alternative behavior, and (e) studies conducted within randomized control trials or concurrent single-case designs with three potential demonstrations of effect. Non-concurrent multiple baseline designs and designs with fewer than three demonstrations of effect (e.g., A–B–A–C designs) were excluded, in that they do not meet contemporary standards for experimental control (Barton et al., 2018; Gast et al., 2018). See Table 1 for complete inclusion criteria.

First, we conducted a systematic search of four electronic databases (PsycINFO, ERIC, PubMed, and ProQuest Dissertations and Theses Global) to identify peer-reviewed and non-peer-reviewed articles examining escape-based interventions using the following terms: *functional analys* and (avoid* or escap* or demand).* For ProQuest Dissertations and Theses Global, the search was limited to titles and abstracts only, to exclude a substantial number of dissertations from unrelated disciplines. We completed this search in January 2019 and updated the search in April 2020.

In April 2019, we conducted a forward and backward reference search of qualifying articles (i.e., we reviewed all sources that cited included studies and sources cited by included studies). For sources listed in the citation index of Web of Science ($n = 28$), we used the Web of Science "cited by" and "citations" features to pool deduplicated citation lists (the forward and backward search, respectively); we then screened listed citations for inclusion. For articles not listed in the citation index ($n = 5$), we pooled deduplicated citation lists using both Google Scholar's and ProQuest's "cited by" features to conduct the forward search, and screened reference lists by hand to conduct the backward search. The ProQuest forward search included the PsycINFO, ERIC, and ProQuest Dissertations and Theses Global databases.

From the sources identified in the forward, backward, and database searches, we evaluated each source against our inclusion criteria, and eliminated articles that did not meet criteria. Through this screening process, we identified 39 articles meeting inclusion criteria, containing 273 study designs. See Fig. 1 for the PRISMA diagram (Moher et al., 2010) outlining screening procedures.

We collected interobserver agreement (IOA) for 100% of screening; that is, each doctoral student independently screened all potential sources, and then, both

Table 1 Inclusion Criteria for Review of Escape-Based interventions Addressing Escape-Maintained Challenging Behavior

Category	Inclusion	Exclusion
Participants	Individuals who engaged in escape-maintained challenging behavior, identified by functional analysis (FA) of challenging behavior; this included individuals with multiply-maintained behavior (i.e., multiple functions, including escape)	Individuals without escape-maintained challenging behavior or with escape function determined without FA of challenging behavior (e.g., descriptive or indirect assessment, FA of alternative behavior)
Intervention	Interventions designed to decrease challenging behavior in which escape was provided contingent on challenging behavior)	Interventions in which escape was withheld contingent on challenging behavior (unless it was compared to an escape-based intervention)
Comparison	Comparisons between escape-based and non-intervention (baseline/control) conditions; comparisons between escaped-based and escape extinction interventions	Comparisons between two or more escape-based interventions
Outcomes	Challenging or alternative behavior measured via observation	Challenging or alternative behavior measured via non-observational methods (e.g., office discipline referrals); other categories of behavior (e.g., response allocation in concurrent operant assessments)
Setting	No setting restrictions	N/A
Study design	Randomized control trials and concurrent single-case/single-subject design with three potential demonstrations of effect, presented in graphic format to allow for visual analysis	Qualitative, correlational, and quasi-experimental design; single-case designs with fewer than three potential demonstrations of effect or without graphic displays; non-concurrent multiple baseline or multiple probe designs

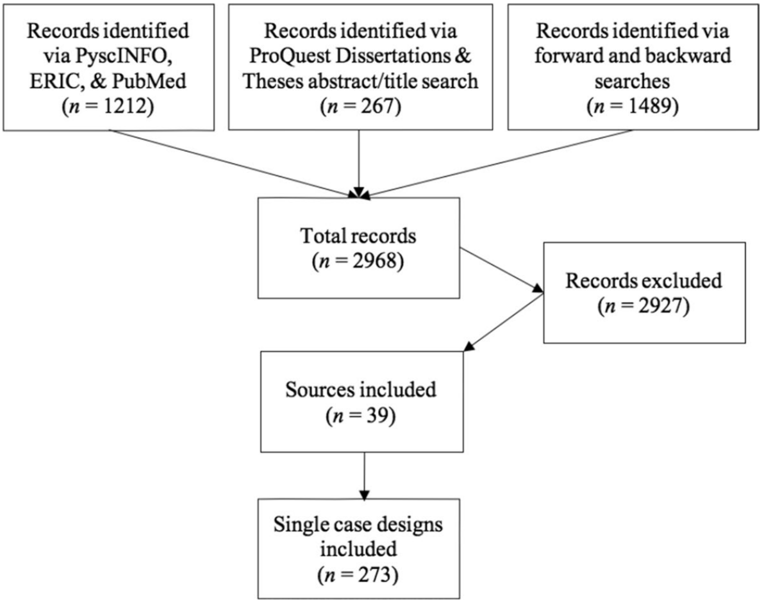

Fig. 1 PRISMA Flow Diagram Outlining Search and Screening Procedures

doctoral students met to compare included sources and to reach consensus on discrepancies. IOA was calculated by dividing the number of include/exclude agreements by the total number of articles screened, then multiplying by 100. IOA was 98.3% overall (98.0% for the database search and 98.7% for the forward/backward search). We conducted discrepancy discussions to reach consensus on all disagreements. If disagreements were not easily resolved, we consulted the third author, a doctoral-level behavior analyst, for a final decision.

Coding

We coded studies on descriptive categories (participants, intervention context, independent variables, and dependent variables) and three outcome categories (quality/rigor, visual analyses, and statistical analyses). Participant information and intervention context were coded separately for each participant and only included participants for whom inclusion criteria were met. Thus, coding may have included one, some, or all participants in a given source. We coded all other variables separately for each treatment design. A design was defined as an experimental comparison with at least three potential demonstrations of effect (Ledford et al., 2018b), and could include a single case (e.g., A–B–A–B design) or multiple cases (e.g., a multiple baseline across participants design). A potential demonstration of effect could be across phases within a withdrawal or multitreatment design, across phases within each tier of a multiple probe or baseline design, or a comparison of two data paths within a single phase in an alternating treatments design. If multiple designs were included in the same graph (i.e., a combination design), each design was analyzed separately (e.g., an A–B–A–B embedded within each tier of a multiple baseline design would be coded as four designs—three A–B–A–B and one multiple baseline design). In multitreatment combination designs, conditions overlapping between comparisons were counted toward both designs, if appropriate (e.g., a B–A–B–A–C–A–C would be coded as two withdrawal designs [B–A–B–A and A–C–A–C]). In accordance with guidance from Ledford and Gast (2018), if a single design included phases that did not contribute to three potential demonstrates of effect, these were excluded from analysis. For example, in an A–B–C–A–C–A design, the final four phases were analyzed and the initial A–B phases were excluded. If qualifying and non-qualifying participants were both included in a multiple baseline or multiple probe design, only qualifying participants were coded for participant and intervention context characteristics. However, in accordance with visual analysis procedures for multiple baseline and probe designs (Gast et al., 2018), all tiers were included when assessing variables at the design level (e.g., assessing presence or absence of functional relation).

We collected IOA for 100% of coding; that is, for every article, each doctoral student independently coded all relevant data, and then, both doctoral students met to compare codes and reach consensus on discrepancies. This included scoring of the *Single-Case Analysis and Review Framework* (SCARF: Ledford et al., 2020). IOA was calculated by dividing the number of agreements by the total possible agreements, and multiplying by 100%. Mean agreement was 95.5% (range 89.7–98.3% by study). By category, average IOA was 95.0% for participants, 95.5% for setting/implementer, 96.4% for independent variables, 97.5% for dependent variables, and 94.8% for SCARF coding. See Supplemental Reliability Data for IOA reporting at the participant, design, study, and category level. We discussed all discrepancies, and consensus was reached for final codes. If disagreements were not easily resolved, we consulted the third author for a decision. See Supplemental Coding Document for the complete codebook with all final decisions.

Participants

We coded the following participant characteristics: age, gender, disability/eligibility categories, and race or ethnicity. We also coded method of FA (e.g., analog, brief, latency-based), topography of challenging behavior (as defined for FA), and functions of challenging behavior assessed via FA. This included coding whether each behavioral function (escape, attention, tangible, automatic/sensory, or other) was tested within the FA and whether there was a functional relation between the test condition and the control or pairwise condition. We used visual analysis to determine the presence of absence of a functional relation, and determined a functional relation to exist (a) when there was little to no overlap between the test and control condition, or (b) when there was some overlap between test and control conditions, but an increasing differentiation over time.

Intervention Context

We coded the setting(s) in which intervention occurred and the role of the implementer(s) who conducted intervention. Setting was coded as home, clinic, school, other, or not reported. Home settings included residential living facilities and group homes. Clinic settings included research laboratories and inpatient hospital units. Implementer was coded as researcher, parent, teacher, staff member, other, or not clearly reported. Codes for paraprofessional, peer, and behavior analyst were also available but never used. To be reported as implementer, the role of the implementer and their relationship to the participant had to be clearly specified. We coded "author" and "we" as researcher; but ambiguous terms such as "therapist" were coded as not clearly reported unless information was provided on the therapist's relationship with the participant.

Independent and Dependent Variables

For independent variables, we coded the design type (e.g., withdrawal, alternating treatments design), comparison type, and intervention components of each design evaluating the escape-based treatment. Comparison types were categorized as comparisons between (a) escape-based and baseline/control conditions (hereafter: baseline comparison), or (b) escape-based and escape extinction conditions (hereafter: escape extinction comparison). Interventions fell into three primary categories: antecedent modifications, consequent manipulations, and packages; see Table 2 for descriptions and examples of each category and subcategory.

Antecedent modifications were environmental changes implemented prior to the occurrence of challenging or alternative behavior. In these interventions, researchers (a) manipulated the environment to make it less aversive, or (b) directly taught skills to learners to increase their ability to advocate for reinforcement (e.g., requesting escape) or cope with aversive stimuli (e.g., diaphragmatic breathing, demand fading). Subcategories of antecedent modifications included non-contingent access to preferred stimuli, pre-session access to preferred stimuli, instructional modifications, diaphragmatic breathing, advanced notice, demand fading, and pre-session skill training (i.e., instruction on the communication response that occurred outside intervention sessions).

We coded consequent manipulations as intervention components that were implemented following alternative behavior. This included differential reinforcement and equal escape. We defined differential reinforcement as providing reinforcement contingent on alternative behavior that is different from or "better than" reinforcement provided on challenging behavior. Subcategories of differential reinforcement included escape to preferred stimuli (i.e., "enriched breaks") and access to preferred stimuli without escape. We also coded equal escape (i.e., reinforcement that is not differential), in which a fixed duration of escape was provided contingent on both alternative and challenging behavior.

Finally, we coded "packages" as interventions that included both antecedent modifications and consequent manipulations. We also coded to confirm that escape was provided contingent on challenging behavior during escape-based interventions

Table 2 Intervention components of escape-based interventions

Intervention component	Definition	Examples
Antecedent modifications	*Environmental changes implemented prior to the occurrence of challenging or alternative behavior*	*See below*
Non-contingent access to preferred stimuli	Provision of preferred stimuli continuously or at pre-specified time intervals during intervention sessions	Continuous access to music (Carey & Halle, 2002), toys (González et al., 2014), earplugs (O'Reilly et al., 2000) or children's videos (Wilder et al., 2005); access to edibles on a variable time schedule (Ingvarsson et al., 2009; Lomas et al., 2010; Lomas Mevers et al., 2014)
Instructional modifications	Changes to instructional procedures to make instruction less aversive	High-probability request sequence (Zarcone et al., 1994a); math number line (Schmidt et al., 2014); math counters (McComas et al., 2000); choice in instructional task order (McComas et al., 2000; Romaniuk et al., 2002)
Pre-session access to preferred stimuli	Provision of preferred stimuli prior to the start of intervention sessions	Pre-session access to sensory activities (Addison et al., 2012), neutralizing routines (Horner et al., 1997), or preferred activities (e.g., listening to music; O'Reilly 2005)
Diaphragmatic breathing	Skill instruction in deep breathing into the diaphragm to increase tolerance to non-preferred stimuli	Take deep breaths to cause object on stomach to rise (Phillips et al., 2019)
Demand fading	Gradually increasing the response requirement prior to receiving access to reinforcement	Gradually increase task provision from 3 per session to 51 per session (Pace et al., 1994); or from 0 per min to 2 per min (Zarcone et al., 1994b)
Advanced notice	Provision of verbal or visual warnings prior to introduction of a non-preferred stimulus	Transition warnings (Vasquez et al., 2017)
Consequent manipulations	*Providing preferred stimuli contingent on alternative behavior*	*See below*
Equal escape for alternative behavior	Provision of equal quality/quantity escape contingent on alternative and challenging behavior	Aggression and compliance each resulted in 30-s break (DeLeon et al., 2001; LaRue et al., 2011; Piazza et al., 1997, 1998)
Differential reinforcement with escape	Provision of equal quality/quantity escape contingent on alternative and challenging behavior; preferred stimuli available only contingent on alternative behavior	Aggression resulted in 30-s break, while compliance resulted in 30-s of access to preferred activities (e.g., sensory activities, computer games; Piazza et al., 1997) or therapist attention (Piazza et al., 1998)

Table 2 (continued)

Intervention component	Definition	Examples
Differential reinforcement with no escape	Preferred stimuli available only contingent on alternative behavior, escape not available contingent on alternative behavior	Aggression resulted in 30-s break, while compliance resulted in access to small edible item and continued task directions (Carter, 2010; DeLeon et al., 2001; Lomas Mevers et al, 2014; Slocum & Vollmer, 2015)
Packages	*Combination of antecedent and consequent manipulations*	*Implementing functional communication training with pre-session skill training* (Adelinis et al., 2001; Kelley et al., 2002; Shirley et al., 1997)

and was *not* provided contingent on challenging behavior during escape extinction interventions.

For dependent variables, we coded the dependent variable category (challenging or alternative behavior) and reporting system (e.g., count/rate, percent of opportunities, percent of intervals, other [with description], or not clearly reported).

SCARF Outcomes

For quality/rigor and visual analysis outcomes, we applied the SCARF (Ledford et al., 2020). The SCARF allows the coder to generate a scatterplot of quality/rigor values and visual analysis values, to gain a clearer picture of the relationship between these two variables in a group of studies.

For strength of quality/rigor, the SCARF allows coders to generate a 0–4 score, with 0 indicating lowest possible quality evidence and 4 indicating highest possible quality evidence. When generating a score for quality/rigor, the SCARF uses a weighted scoring system that weighs rigor components (i.e., interobserver agreement, procedural fidelity, and sufficiency of data collection) twice as heavily as quality components (i.e., social and ecological validity, descriptions of study components). While the primary outcome scoring only allows for whole number scores (i.e., 0, 1, 2, 3, 4, and in our case, − 1), the weighted scoring system used for quality/rigor allows for scores to fall between whole numbers (e.g., 1.7, 2.4).

For primary outcomes, the SCARF allows coders to use visual analysis to generate a 0–4 score, which indicates strength of functional relation. For this measure, scores of 0 indicate consistent equivocal effects, scores of 1 indicate fewer than three demonstrations of effect, scores of 2 indicate at least three demonstrations of effect and one or more equivocal effects, scores of 3 indicate at least three demonstrations of effect (one or more of which are weak) with no equivocal effects, and scores of 4 indicate at least three moderate or strong demonstrations of effect with no equivocal effects (Ledford et al., 2020). Determination of equivocal, weak, and strong effect was made via visual analysis. For example, in assessing data paths between a baseline and intervention condition within a withdrawal design, we would qualify (a) an equivocal effect as substantial overlap between data paths with no clear level difference, (b) a weak effect as a clear difference in level between data paths, with some overlap or small magnitude change, and (c) a strong effect as clear and consistent difference in level between data paths, with minimal or no overlap and a large magnitude change. We adapted the visual analysis scores to add a possible − 1 to indicate functional relations in favor of the baseline or escape extinction condition.

We also assessed presence or absence of a functional relation (Ledford et al., 2018c). If a functional relation was present, we reported which intervention was more effective (e.g., that the escape-based intervention was more effective than the baseline condition). If no functional relation was present, we reported that results were equivocal. That is, both the escape-based intervention and comparison condition (i.e., baseline condition or escape extinction intervention) resulted in similar behavior patterns. For comparisons of escape-based and escape extinction interventions, equivocal results do not describe the therapeutic results of either intervention;

that is, both interventions may have been equally effective or equally ineffective in reducing challenging behavior and/or increasing adaptive behavior.

To increase validity of the results (i.e., increase confidence that results are not the result of quality/rigor issues), we limited visual and statistical outcomes to designs with SCARF quality/rigor scores of 1.7 or higher. Although that score is somewhat arbitrary, having a minimum acceptable value allows researchers to avoid interpreting outcomes of studies that are most likely to have threats to internal validity that may render outcomes unbelievable. Previous research (e.g., Chazin et al., 2021; Ledford & Windsor, 2021) has included similar rules.

Statistical Outcomes

Baseline conditions were homogeneous, which made comparisons between baseline conditions and escape-based interventions appropriate to synthesize. Typically, baseline conditions were similar to escape conditions in functional analyses, in that implementers provided escape contingent on challenging behavior, and appropriate behaviors did not result in escape (e.g., manding was often ignored, compliance often resulted in brief praise and presentation of the next demand). Comparisons across escape extinction comparisons were heterogeneous (e.g., several studies compared sensory-based treatments to escape extinction interventions while others compared attention extinction interventions to escape extinction interventions). Because magnitude differences between widely varying escape-based interventions and escape extinction conditions would not be expected to be the same, a single synthesis metric was not meaningful; as such, we did not meta-analyze these data.

To estimate the magnitude of effect for baseline comparisons, we calculated log response ratios (LRRs) for each adjacent comparison within a design (Pustejovsky, 2018) and combined estimates using meta-analysis. LRR describes the proportional change of behavior between two conditions. We used LRR decreasing (LRRd) to estimate the magnitude of effect on challenging behavior (where a therapeutic effect was a decrease in the dependent variable); we used LRR increasing (LRRi) to estimate the magnitude of effect on alternative behavior (where a therapeutic effect was an increase in the dependent variable).

To calculate LRR, we first digitized data points from publication graphs using PlotDigitizer (Huwaldt & Steinhort, 2013). Coders overlapped on 16.2% of graphs. Data point values were considered to be in agreement when the difference between values was equal to or less than 2% of the maximum y value indicated on the graph. IOA was 98.9% overall (range: 87.5–100%). Second, we calculated LRR for each A–B comparison in each design (e.g., an A–B–A–B withdrawal contained two comparisons while an alternating treatments design with two data paths contained one) using a web-based calculator (Pustejovsky & Swan, 2018). Third, we combined LRR values using a multilevel random effects meta-analysis model using the metafor (Viechtbauer, 2010) package in R statistical environment (R Core Team, 2016). To estimate the overall effect of baseline comparisons, we used a cluster-robust variance estimation with the clubSandwich package (Pustejovsky, 2019).

LRR cannot be calculated if the mean of a condition is at floor or ceiling levels (e.g., all data within a condition are 0 or 100%) or if a condition contains a single datum point. Overall, 17 designs were excluded due to floor means, 1 due to ceiling means, and 1 due to a combination of floor means and a single datum point. Additionally, three designs with latency-based measures were excluded, because data from these designs are not comparable to designs measuring the occurrence (e.g., count, duration) of behavior.

Results

We identified 39 studies (38 peer-reviewed articles and 1 dissertation) that met inclusion criteria, containing 273 study designs. These designs included 158 baseline comparisons and 115 escape extinction comparisons. Withdrawal designs (e.g., A–B–A–B; $n=156$; 57.1%), multitreatment designs (e.g., A–B–C–B–C; $n=66$; 24.2%), and alternating treatments designs ($n=47$; 17.2%) were the predominant design types, along with 4 multiple baseline designs across participants (1.5%).

Participants

Studies included 79 participants, 60 males (75.9%) and 19 females (24.1%). Participant ages fell between 1 and 49 years (median: 8). Twenty-nine participants were young children (1–6 years old; 36.7%), 37 were school-aged (7–12 years old; 46.8%), 7 were adolescents (13–19 years old; 8.9%), and 6 were adults (27–49 years old; 7.6%). The majority were diagnosed with autism spectrum disorders ($n=30$; 38.0%), intellectual or developmental disabilities ($n=29$; 36.7%), multiple/severe disabilities ($n=16$; 20.2%), or were identified as being at risk for disabilities ($n=15$; 19.0%). Ten participants were diagnosed with emotional and behavior disorders (12.7%); 8 with attention-deficit/hyperactivity disorder (10.1%), 7 with developmental delay (8.9%), and 6 with seizure disorders (7.6%). The sum of these values exceeds 100%, because many participants ($n=33$; 41.8%) had comorbid diagnoses. Five or fewer participants met criteria for remaining diagnoses: vision and/or hearing impairment ($n=5$; 6.3%), speech/language impairment ($n=3$; 3.8%), feeding disorder ($n=3$; 3.8%), traumatic brain injury ($n=2$; 2.5%), and specific learning disability ($n=1$; 1.3%). Race and ethnicity were not reported for the majority of participants ($n=76$; 96.2%); all three that were reported were Caucasian (3.8%).

Behavioral functions were commonly identified via analog FAs ($n=40$ participants; 50.6%). Remaining functions were identified via pairwise FAs ($n=23$; 29.1%), brief FAs ($n=5$; 6.3%), multiple FA methods ($n=8$; 10.1%), a trial-based FA ($n=1$; 1.3%), or a simplified traditional FA ($n=1$; 1.3%). The FA for one participant was not described sufficiently to categorize. For most participants ($n=44$; 55.7%), escape was the only identified function of behavior. For 26 participants, challenging behavior was multiply-maintained. That is, behavior was maintained by escape, plus access to attention ($n=9$; 11.4%), tangibles ($n=9$; 11.4%), automatic reinforcement ($n=1$; 1.3%), or a combination of these ($n=7$; 8.9%). For remaining participants ($n=9$; 11.4%), FA results were not graphed to allow for visual analysis.

Nearly all participants ($n = 59$; 74.7%) engaged in at least one form of dangerous behavior (i.e., those likely to cause harm to self or others), including physical aggression ($n = 37$; 46.8%), self-injurious behavior ($n = 10$; 12.7%), or both ($n = 12$; 15.2%). Participants also engaged in property destruction ($n = 33$; 41.8%), non-compliance ($n = 26$; 32.9%), inappropriate verbalizations ($n = 20$; 25.3%), and/or elopement ($n = 5$; 6.3%). The sum of these values exceeds 100%, because the majority of participants ($n = 46$; 58.2%) engaged in multiple topographies of challenging behavior. Additionally, five participants (6.3%) had additional topographies of behavior that did not meet criteria for these categories, including tantrums ($n = 3$; 3.8%), hand wringing ($n = 1$; 1.3%), and head shaking ($n = 3$; 1.3%).

Intervention Context

Most participants took part in studies conducted in clinics ($n = 46$; 58.2%). Fewer participated in studies conducted in authentic settings (schools, $n = 17$; 21.5%; homes, $n = 7$; 8.9%). For two participants (2.5%), studies were conducted in both school and clinic. For one participant (1.3%), the study was conducted in a nurse's office. For six participants (7.6%), setting was not reported with sufficient detail for categorization.

More than one-third of implementers ($n = 32$; 40.5%) were described with insufficient detail for categorization. Other implementers included researchers (e.g., study authors, graduate students; $n = 27$; 34.2%), teachers ($n = 10$; 12.7%), parents ($n = 5$; 6.3%), or staff members ($n = 2$; 2.5%). For three participants (3.8%), interventions were conducted by both a researcher and parent.

Independent Variables

Across all escape-based interventions, implementers reinforced every instance of challenging behavior with access to escape, typically for the same duration provided contingent on challenging behavior in baseline conditions. One hundred designs included exclusively antecedent modifications. Antecedent modifications included use of: non-contingent access to preferred stimuli ($n = 61$; 22.3%), instructional modifications ($n = 20$; 7.3%), pre-session access to preferred stimuli ($n = 12$; 4.4%), diaphragmatic breathing ($n = 3$; 1.1%), advanced notice ($n = 2$; 0.7%), and demand fading ($n = 1$; 0.4%). One intervention combined demand fading with non-contingent access to social interaction. There were two types of consequent manipulations: differential reinforcement ($n = 79$; 28.9%) and equal escape ($n = 70$; 25.6%). Differential reinforcement included access to preferred stimuli with ($n = 56$; 20.5%) or without ($n = 23$; 8.4%) escape from aversive stimuli contingent on alternative behavior. Finally, 24 interventions (8.8%) included a combination of antecedent and consequent manipulations. All packages were variations of functional communication training, and included pre-session skill training as the antecedent modification. Most packages ($n = 17$; 6.2%) included differential reinforcement, and remaining interventions ($n = 7$; 2.6%) included equal escape.

Dependent Variables

Across designs, interventions targeted reducing challenging behaviors (e.g., self-injurious behavior, aggression, elopement; $n = 162$; 59.3%) or increasing alternative behaviors (e.g., requesting, compliance, food acceptance; $n = 111$; 40.7%). Challenging behaviors were primarily reported as a number of occurrences or estimate of number: count/rate ($n = 108$; 39.6%), percentage of intervals (using partial interval time sampling; $n = 14$; 5.1%), or percentage of opportunities ($n = 10$; 3.7%). In remaining designs, challenging behaviors were reported as a percentage of session time ($n = 30$; 11.0%; i.e., duration).

Alternative behaviors were primarily reported as a number of occurrences or estimate of number: percentage of opportunities ($n = 94$; 34.4%), count/rate ($n = 6$; 2.2%), or percentage of intervals (using partial interval time sampling; $n = 6$; 2.2%). Remaining variables were reported as latency to alternative behaviors ($n = 3$; 1.1%) or percentage of session time ($n = 1$; 0.4%). For one dependent variable, the measurement system was not described.

Outcomes

Of the 273 total designs, we assessed quality/rigor and visual analysis outcomes for 222 designs (81.3%) with quality/rigor scores of 1.7 or higher. Including only high-quality designs increases likelihood that analyzed outcomes reflect functional relations between independent and dependent variables, rather than threats to internal validity. The outcomes analysis included 120 baseline comparisons (54.1%), with 69 designs (31.1%) assessing challenging behavior and 51 designs (23.0%) assessing alternative behavior. This also included 103 escape extinction comparisons (46.4%), with 66 designs (29.7%) assessing challenging behavior and 37 designs (16.7%) assessing alternative behavior. We conducted statistical analysis outcomes for baseline comparisons only, given that purposes across escape extinction comparisons were heterogeneous and thus inappropriate to synthesize.

Baseline Comparison Outcomes

Quality/Rigor Outcomes SCARF quality/rigor scores for baseline comparisons are plotted on the *x*-axis of Fig. 2 (top panel). Across the full set of designs, scores ranged from 0.8 to 2.7 (mean: 1.8). For the set of 120 designs with SCARF quality/rigor scores of 1.7 or higher, scores ranged from 1.7 to 2.7 (mean: 2.0). Lack of procedural fidelity data threatened internal validity across high-quality studies—only 40.0% of designs ($n = 48$) reported fidelity data. Further, only 8.3% of designs ($n = 10$) reported social validity data, and only 34.2% ($n = 41$) utilized authentic settings or implementers. Rigor was strengthened by adequate collection of data ($n = 117$; 97.5%) and IOA ($n = 120$; 100%).

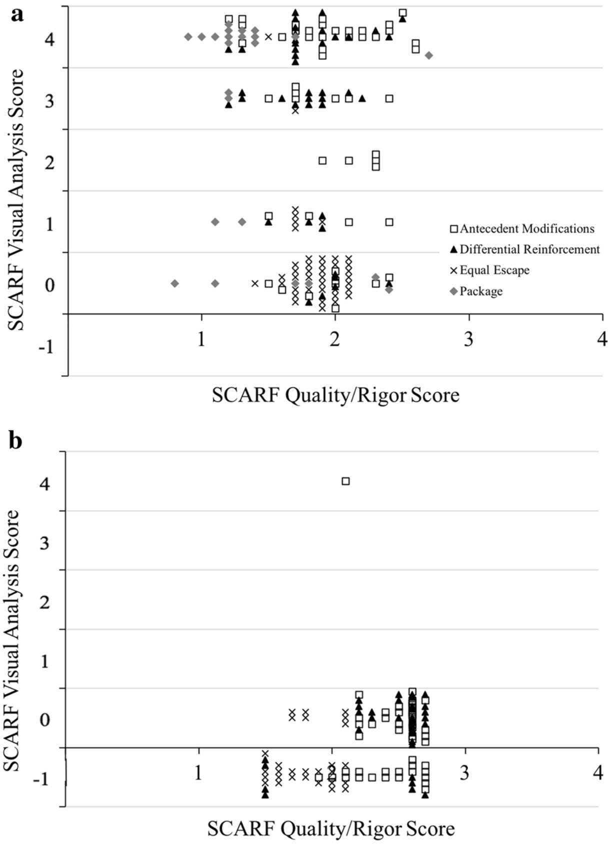

Fig. 2 SCARF Scatterplots for Baseline Comparisons (Top Panel) and Escape Extinction Comparisons (Bottom Panel). *Note:* Scatterplot of scores for SCARF quality/rigor (continuous scale: 0 to 4) and visual analysis outcomes (categorical scale: – 1 to 4) for baseline comparisons (top panel) and escape extinction comparisons (bottom panel). Scores of – 1 indicate that comparison conditions (i.e., baseline, escape extinction) that were more effective than escape-based interventions

Visual Analysis Outcomes SCARF visual analysis scores for baseline comparisons are plotted on the *y*-axis of Fig. 2 (top panel). Across the full set of designs, scores ranged from – 1 to 4 (median: 3). For the set of 120 designs with SCARF quality/rigor scores of 1.7 or higher, scores also ranged from – 1 to 4 (median: 1.5). Challenging behavior was measured in 68 of these designs. Of these, 35 showed that escape-based interventions were more effective than baseline conditions (51.5%), 32 reported equivocal effects (47.1%), and 1 showed that the baseline condition was more effective than the escape-based intervention (1.5%). Alternative behavior was measured in 52 of these designs. Of these, 22 reported that escape-based interventions were more effective than baseline conditions (42.3%) and 30 reported equivocal effects (57.7%).

Table 3 Intervention Components of Escape-Based Interventions for High-Quality Studies*

Intervention component	Baseline comparisons *n* (%)				Escape extinction comparisons *n* (%)			
	Total designs	Escape-based interventions more effective than baseline conditions	Equivocal effects	Baseline conditions more effective than escape-based interventions	**Total designs**	Escape-based interventions more effective than escape extinction interventions	Equivocal effects	Escape extinction interventions more effective than escape-based interventions
Antecedent modifications	**40**	27 (67.5%)	13 (32.5%)	0 (0.0%)	**50**	1 (2.0%)	30 (60.0%)	19 (38.0%)
Non-contingent access to preferred stimuli	**24**	12 (50.0%)	12 (50.0%)	0 (0.0)	**33**	0 (0.0%)	24 (72.7%)	9 (27.3%)
Instructional modifications	**8**	7 (87.5%)	1 (12.5%)	0 (0.0%)	**11**	1 (9.1%)	6 (54.5%)	5 (45.5%)
Pre-session access to preferred stimuli	**3**	3 (100.0%)	0 (0.0%)	0 (0.0)	**6**	0 (0.0%)	0 (0.0%)	6 (100.0%)
Diaphragmatic breathing	**3**	3 (100.0%)	0 (0.0%)	0 (0.0%)	**0**	N/A	N/A	N/A
Demand fading	**2**	2 (100.0%)	0 (0.0%)	0 (0.0)	**0**	N/A	N/A	N/A
Advanced notice	**2**	2 (100.0%)	0 (0.0%)	0 (0.0%)	**0**	N/A	N/A	N/A
Consequent manipulations	**73**	27 (37.0%)	46 (73.0%)	0 (0.0)	**53**	0 (0.0%)	36 (67.9%)	17 (32.1%)
Equal escape for alternative behavior	**38**	1 (2.6%)	36 (97.4%)	1 (2.6%)	**20**	0 (0.0%)	7 (35%)	13 (65.0%)
Differential reinforcement with escape	**14**	12 (85.7%)	2 (14.3%)	0 (0.0)	**33**	0 (0.0%)	29 (87.9%)	4 (12.1%)

Table 3 (continued)

Intervention component	Baseline comparisons *n* (%)				Escape extinction comparisons *n* (%)			
	Total designs	Escape-based interventions more effective than baseline conditions	Equivocal effects	Baseline conditions more effective than escape-based interventions	**Total designs**	Escape-based interventions more effective than escape extinction interventions	Equivocal effects	Escape extinction interventions more effective than escape-based interventions
Differential reinforcement with no escape	**21**	14 (66.7%)	7 (33.3%)	0 (0.0%)	**0**	N/A	N/A	N/A
Packages	**6**	2 (33.3%)	4 (66.7%)	0 (0.0%)	**0**	N/A	N/A	N/A

*Results indicated only for high-quality studies (i.e., studies with SCARF quality/rigor scores at or above 1.7)

Component Efficacy in High-Quality Studies See Table 3 for the incidence of functional relations for each intervention component within high-quality studies. Antecedent modifications were effective in 27 of 40 designs (69.2%). Most of these studies ($n = 24$) assessed non-contingent access to preferred stimuli, and these were effective for 50% of designs. Instructional modifications were effective in 7 of 8 designs (87.5%). Remaining antecedent modifications were effective for all designs, and included pre-session access to preferred stimuli, diaphragmatic breathing, demand fading, and advanced notice. However, it should be noted that there are relatively few designs (2–4) per intervention type.

Differential reinforcement interventions were effective in 26 of 35 designs (74.3%). For designs in which alternative behavior resulted in escape to preferred stimuli (i.e., "enriched breaks"), 85.7% of designs indicated a functional relation. For designs in which alternative behavior resulted in access to preferred stimuli (without escape), 66.7% of designs indicated a functional relation. Equal escape interventions were almost never effective, with 1 of 37 designs (2.6%) indicating that the escape-based intervention was more effective than the baseline condition and 1 design (2.6%) indicating that the baseline condition was more effective than the intervention condition.

Finally, six packaged interventions assessed variations of functional communication training (FCT), combining pre-session skill training with differential reinforcement or equal escape. Similar to results above, FCT with differential reinforcement tended to be effective ($n = 2$ of 3 indicating a functional relation), while FCT with equal escape was ineffective ($n = 0$ of 3 indicating a functional relation).

Statistical Outcomes We calculated 67 LRRd estimates from 23 studies for baseline comparisons (see Fig. 3); 23 designs were dropped due to low SCARF scores, 6 due to floor means, 1 due to ceiling means, and 1 due to a combination of floor means and one datum point. Escape-based interventions significantly decreased challenging behavior, LRRd $= -1.02$, $SE = 0.14$, 95% CI: -1.25, -0.79]. We calculated 37 LRRi estimates from 15 studies; 15 designs were dropped due to low SCARF scores, 11 due to floor means, and 3 due to latency measures. Escape-based interventions significantly increased alternative behavior, LRRi $= 0.56$, $SE = 0.12$, 95% CI: [0.33, 0.79].

EE Comparison Outcomes

Quality/Rigor Outcomes SCARF quality/rigor scores for escape extinction comparisons are plotted on the *x*-axis of Fig. 2 (bottom panel). Across the full set of designs, scores ranged from 1.5 to 3.0 (mean: 2.4). For the set of 103 designs with SCARF quality/rigor scores of 1.7 or higher, scores ranged from 1.7 to 2.7 (mean: 2.4). Of the high-quality designs, none included social validity measures and only 18.4% included authentic settings or implementers. Conversely, authors reported sufficient IOA (100%), data per condition (100%), and procedural fidelity (73.8%).

Visual Analysis Outcomes SCARF visual analysis scores for escape extinction comparisons are plotted on the *y*-axis of Fig. 2 (bottom panel). For these, -1 indicated

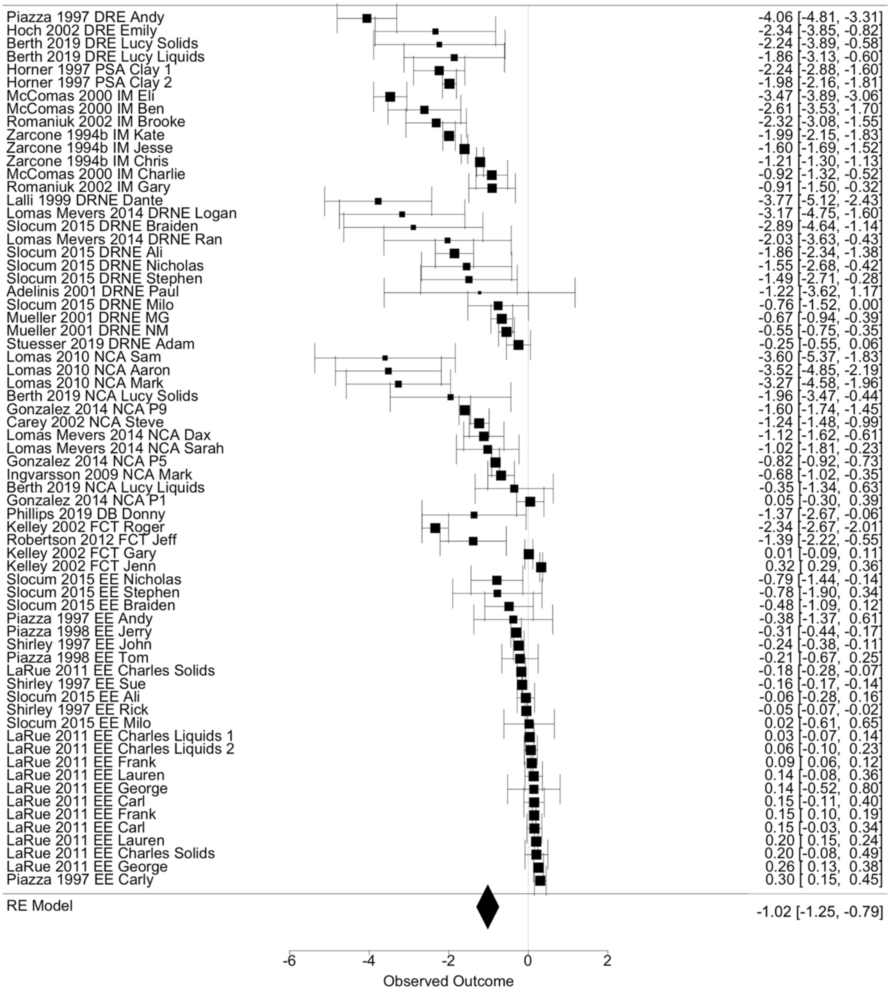

Fig. 3 Forest Plots for Baseline Comparisons Measuring Challenging Behavior (Top Panel) and Alternative Behavior (Bottom Panel). *Note:* For analyses of log response ratio (LRR) decreasing (top panel), negative effect size indicates a therapeutic effect. For analyses of LRR increasing (bottom panel), positive effect size indicates a therapeutic effect. Designs are organized by intervention type. DRE = differential reinforcement with escape; PSA = pre-session access to preferred stimuli; IM = instructional modifications; DRNE = differential reinforcement with no escape; NCA = non-contingent access to preferred stimuli; DB = diaphragmatic breathing; FCT = functional communication training; EE = equal escape

that escape extinction interventions were more effective than escape-based interventions, and 0 represented equivocal effects. Across the full set of designs, scores ranged from − 1 to 4, although all but one fell between − 1 and 0 (median: 0). For the set of 103 designs with SCARF quality/rigor scores of 1.7 or higher, scores ranged from − 1 to 4 (median: 0). Challenging behavior was measured in 66 of these designs. Of these, 46 reported equivocal effects (69.7%), 19 showed that escape extinction interventions were more effective than escape-based interventions; 28.8%), and 1 showed that the

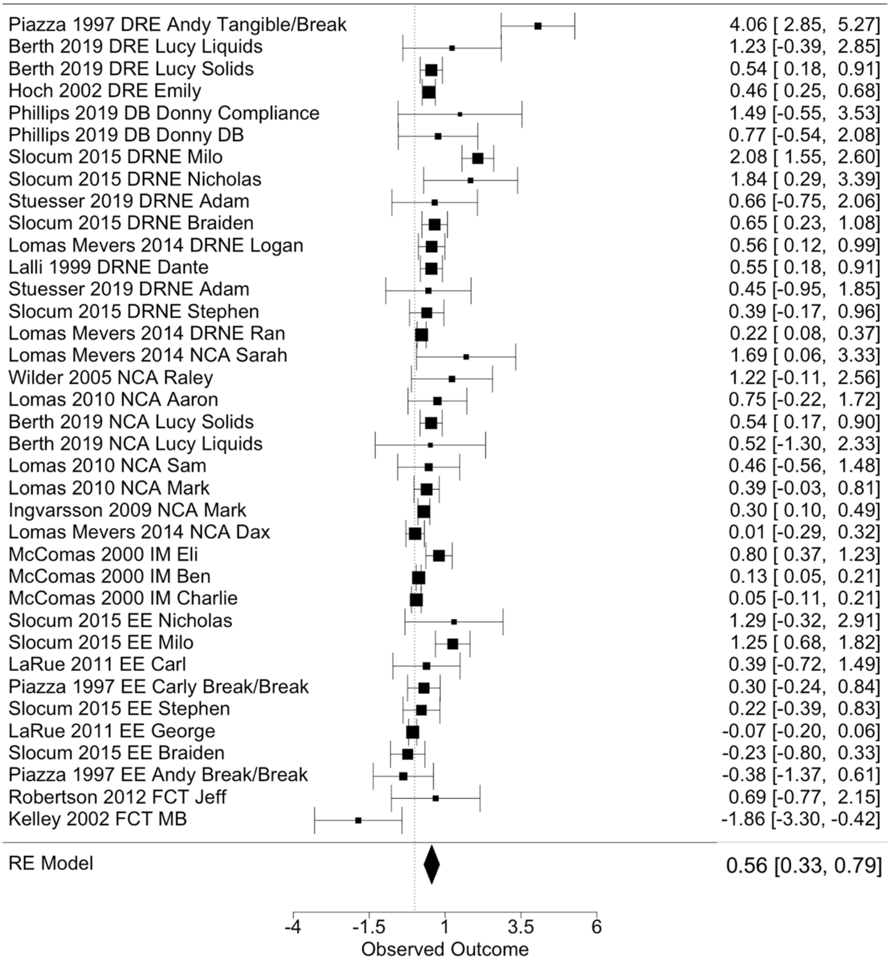

Fig. 3 (continued)

escape-based intervention was more effective than the escape extinction intervention (1.5%). Alternative behavior was measured in 37 of these designs. Of these, 20 reported equivocal effects (54.1%) and 17 indicated that escape extinction interventions were more effective than escape-based interventions (45.9%).

Component Efficacy in High-Quality Studies See Table 3 for the incidence of functional relations for each intervention component within high-quality studies. For comparisons in which the escape-based intervention included antecedent modifications, 31 of 50 designs (62%) reported equivocal effects. Most of these studies ($n=33$; 66%) assessed non-contingent access to preferred stimuli; these reported equivocal effects for 72.7% of designs. Of 11 designs assessing instructional modifications, 6 reported equivocal effects (54.5%), 4 reported that escape extinction interventions were more effective than escape-based interventions (36.4%), and 1 reported that an escape-based intervention was more effective than an escape extinction intervention (9.1%). Of 6 designs assessing pre-session access to preferred stimuli, all reported that escape extinction interventions were more effective than escape-based interventions (100%); in these studies, participants were provided pre-session access to sensory integration activities, though a sensory condition was not included in the participants' functional analyses (Addison et al., 2012).

For comparisons in which the escape-based intervention included differential reinforcement, equivocal effects were reported for 29 of 33 designs (87.9%). For all of these designs, alternative behavior resulted in escape to preferred stimuli (i.e., "enriched breaks").

For comparisons in which the escape-based intervention included equal escape, 13 of 20 designs (65.0%) reported that escape extinction interventions were more effective than escape-based interventions. No escape extinction comparisons included escape-based packaged interventions.

Discussion

The primary purposes of this review were to determine (a) intervention components and contexts in which escape-based interventions have been assessed, (b) rigor and quality of this evidence base, and (c) the efficacy of escape-based interventions, compared to baseline and escape extinction conditions. To our knowledge, this is the first review to synthesize escape-based interventions. Additionally, we were only able to identify two closely related literature reviews: Dart et al. (2018) reviewed interventions for escape-maintained challenging behavior in school settings, and Trump et al. (2019) reviewed differential reinforcement without extinction. This review makes several novel contributions to the literature base. First, our search was more inclusive, screening in all escape-based interventions without escape extinction (cf. Trump et al., 2019) used in all settings with learners of any age or ability (cf. Dart et al., 2018). Second, we included unpublished dissertations and theses, to counteract publication bias (Ledford et al., 2018a; Shadish et al., 2016). Third, we ensured that all sources, including non-peer-reviewed sources, were not included in outcomes analyses if they did not meet contemporary rigor standards (Ledford et al., 2018b). This reduced the likelihood that estimates were impacted by studies whose findings were attributable to variables outside the intervention (e.g., threats to internal validity). Finally, we utilized meta-analysis procedures to quantitatively examine outcomes for escape-based interventions.

Antecedent modifications were more effective than baseline conditions in decreasing challenging behavior and increasing alternative behavior (see Table 2). These interventions work by manipulating motivating operations to decrease the value of escape. Despite the efficacy of these interventions, only 11 high-quality baseline comparison designs assessing antecedent modifications incorporated some form of skill instruction (e.g., pre-session skill training, tolerating aversive stimuli via demand fading). These critical skills may lead to sustained behavior change because they provide a long-term, functional alternative to challenging behavior likely to be reinforced in typical contexts. Other antecedent-based interventions may be less likely to result in maintained outcomes when intervention is withdrawn or poorly implemented and may not generalize to contexts without intervention present. For example, interventions that are context dependent (e.g., use of non-contingent reinforcement or high-probability request sequences) may lead to temporary decreases in escape-maintained challenging behavior, while teaching skills that can persist outside of intervention (e.g., teaching tolerance with demand fading, teaching emotional regulation with diaphragmatic breathing) may be more likely to lead to long term, generalized changes.

Access to preferred stimuli (with and without escape) contingent on alternative behavior was more effective than baseline conditions in affecting therapeutic outcomes. These interventions capitalize on the value of varying quality and quantity of reinforcement, and may be most effective when matched to the function(s) of the learner's behavior and use of their most potent reinforcers. Equal escape interventions were not effective compared to baseline conditions; these results are consistent with matching law theory (Borrero et al., 2010; Catania, 2013), given that equal escape does not capitalize on reinforcement value and learners typically have long histories of reinforcement with challenging behavior. Thus, it may be most effective for practitioners to differentiate contingencies of reinforcement for alternative and challenging behavior, ensuring that the learner prefers reinforcement provided contingent on alternative behavior more than reinforcement provided contingent on challenging behavior. Given that "enriched breaks" were the most effective consequent strategy assessed, practitioners may consider providing breaks contingent on challenging behavior, as well as providing breaks with preferred items contingent on alternative behavior. For individuals with multiply-maintained challenging behavior, "enriched breaks" should include access to stimuli shown to previously maintain challenging behavior (e.g., breaks to preferred forms of attention for challenging behavior maintained by access to escape and attention).

The majority of escape extinction comparisons reported equivocal effects, in which escape extinction interventions were not more effective than escape-based interventions. Interestingly, intervention components that were effective in baseline comparisons (e.g., instructional modifications, differential reinforcement with escape) most often reported equivocal effects in escape extinction comparisons. Similarly, intervention components that reported a higher proportion of equivocal effects in baseline comparisons (e.g., equal escape, non-contingent access to preferred stimuli) also reported a higher proportion of designs in which escape extinction interventions was more effective than escape-based interventions. Pre-session access was the exception to this rule; however, the pre-session stimuli matched the function of the learner's behavior in baseline comparisons, but not in escape extinction comparisons. Thus, functional match may moderate efficacy of pre-session access. Overall, these results indicate that if escape-based interventions include components that are effective in baseline comparisons (e.g., instructional modification, "enriched breaks"), escape-based interventions may be similarly as effective as escape extinction interventions. Thus, if practitioners opt to use escape-based interventions, they may find it most useful to use instructional modifications and differential reinforcement with escape (i.e., “enriched breaks”), as these were most likely to be reported effective in baseline comparisons and least likely to be less effective than escape extinction interventions.

Limitations

Reporting issues limit confidence with which we can interpret outcomes. Because procedural fidelity data were reported for relatively few designs, particularly for baseline comparisons, there is decreased confidence that baseline and intervention components were implemented as described. Further, the majority of studies did not describe the implementer; thus, it is difficult to know what training and experience are needed to effectively implement these interventions.

Limitations of the research base also decrease our ability to determine for whom and under what contexts escape-based interventions are effective. Across most comparisons, lack of social validity measures, authentic implementers (parents, teachers), and typical settings (homes, schools) make it difficult to assess whether escape-based interventions can be conducted in usual contexts by indigenous implementers and whether practitioners are willing to implement them.

Although this review included a number of interventions intended to increase food acceptance for individuals with feeding disorders and related issues (e.g., Addison et al., 2012; Berth et al., 2019; LaRue et al, 2011; Wilder et al., 2005), there were fewer feeding studies than may have been expected, given that food refusal is typically assumed to be escape-maintained (Ledford et al., 2018a). However, very few feeding interventions utilize FAs to confirm an escape function, and those that have confirmed escape functions via FA historically included escape extinction in intervention (Ledford et al., 2018a). As the first review to assess escape-based interventions broadly, we thought it important to use stringent inclusion criteria confirming an escape function experimentally, as results of descriptive functional assessments (e.g., interviews, direct observations) tend to have low consistency with experimental analyses (Alter et al., 2008). Future reviews on escape-based interventions might expand inclusion criteria to include descriptive functional assessments, in order to assess a wider pool of escape-based interventions, particularly to address food refusal.

This review was limited to studies that explicitly provided reinforcement contingent on challenging behavior. Thus, we excluded other strategies that may include escape extinction while mitigating its side effects, including (a) the use of partial escape extinction (i.e., escape extinction that excludes full physical guidance) and (b) providing escape contingent on low effort, alternative responses on a continuous schedule of reinforcement in conjunction with demand fading, such that response effort continues to increase, while access to escape remains consistent. For example, the enhanced choice model for skill-based treatment incorporates escape extinction in that challenging behavior results in continued provision of task directions (Rajaraman et al., 2021). However, this model mitigates side effects of escape extinction by (a) using partial escape extinction (paced verbal model prompting) rather than full escape extinction (full physical prompting) and (b) in conjunction with demand fading, making escape continuously available contingent on a low effort alternative response (e.g., walking to the "hangout" portion of the room), but not contingent on challenging behavior. To further understand strategies for mitigating the escape of escape extinction, future researchers might consider synthesizing evidence for these alternative intervention types.

A final limitation of this study is lack of empirical estimates for single-case effect sizes (i.e., what effects are likely for these populations, behaviors, and contexts). Because few other single-case meta-analyses have used LRR to date, it is difficult to determine how efficacy of these interventions compares to those with similar dependent variables. Further, LRR effect size metrics cannot be compared to those calculated for group comparison studies.

Implications for Research

First, additional research conducted in authentic contexts is needed. These studies should include measures of procedural fidelity and social validity for authentic implementers, to assess whether they can and will implement escape-based interventions with fidelity. Second, research is needed to inform recommendations to practitioners regarding least restrictive interventions: behavior analysts, teachers, and school psychologists need transparent guidelines on weighing the costs and benefits of interventions that decrease restrictiveness (escape-based interventions) versus those that may be more efficient (escape extinction interventions). Future studies (e.g., group designs) might compare efficiency of these two intervention types. Finally, researchers should assess strategies designed to concurrently maximize efficiency and minimize side effects. Interventions including complements to escape extinction (e.g., enhanced choice model; Rajaraman et al., 2021) may be one solution, but limited peer-reviewed data exist.

Recommendations for Practice

Given the findings of this review, we recommend three considerations for practitioners. First, escape-based interventions that use antecedent modifications and differential reinforcement are largely effective for reducing challenging behavior and increasing alternative behavior, compared to baseline conditions. They can be considered in lieu of escape extinction, in order to mitigate side effects of escape extinction. If progress monitoring indicates insufficient behavior change, assess whether escape-based interventions are conceptually systematic and implemented with high fidelity before adding escape extinction.

Second, regardless of use of escape extinction, use antecedent modifications to decrease the value of escape as a reinforcer and teach alternative behavior. Effective antecedent modifications include use of instructional modifications (including choice), advanced notice, and training to request escape and tolerate the presence of aversive stimuli.

Third and finally, regardless of use of escape extinction, systematically teach and differentially reinforce alternative behavior to increase likelihood of appropriate responding. Contingent on alternative behavior (e.g., requesting, compliance), provide escape to preferred tangibles and attention.

Supplementary Information

Supplemental materials for this article are available on the Open Science Framework at https://osf.io/mf3as/

Acknowledgements We would like to thank Marney S. Pollack, Dr. Johanna L. Staubitz, and Dr. Gregory P. Hanley for their thoughtful feedback on early iterations of this project.

Author Contributions All authors contributed to the conception of this review. The literature search and data analysis were performed by Kate Chazin and Marina Velez. The first draft of the manuscript was written by Kate Chazin and Marina Velez. The work was critically revised by Jennifer Ledford. All authors read and approved the final manuscript.

Declarations

Conflict of interest As a systematic review with no human or animal participants, ethics approval from an institutional review board was not required. There are no conflicts of interest to disclose. No financial support has been provided for the research, authorship, or publication of this article. No materials in this manuscript are the copyrighted work of another individual or organization.

References

*References included in the literature review

*Addison, L. R., Piazza, C. C., Patel, M. R., Bachmeyer, M. H., Rivas, K. M., Milnes, S. M., & Oddo, J. (2012). A comparison of sensory integrative and behavioral therapies as treatment for pediatric feeding disorders. *Journal of Applied Behavior Analysis, 45*, 455–471.

*Adelinis, J. D., Piazza, C. C., & Goh, H. (2001). Treatment of multiply controlled destructive behavior with food reinforcement. *Journal of Applied Behavior Analysis, 34*, 97–100.

Alter, P. J., Conroy, M. A., Mancil, G. R., & Haydon, T. (2008). A comparison of functional behavior assessment methodologies with young children: Descriptive methods and functional analysis. *Journal of Behavioral Education, 17*, 200–219.

Barton, E. E., Lloyd, B. P., Spriggs, A. D., & Gast, D. L. (2018). Visual analysis of graphic data. In J. R. Ledford & D. L. Gast (Eds.), *Single case research methodology* (pp. 239–281). Routledge.

Beavers, G. A., Iwata, B. A., & Lerman, D. C. (2013). Thirty years of research on the functional analysis of problem behavior. *Journal of Applied Behavior Analysis, 46*, 1–21.

*Berth, D. P., Bachmeyer, M. H., Kirkwood, C. A., Mauzy, C. R., IV., Retzlaff, B. J., & Gibson, A. L. (2019). Noncontingent and differential reinforcement in the treatment of pediatric feeding problems. *Journal of Applied Behavior Analysis, 52*, 622–641.

Borrero, C. S., Vollmer, T. R., Borrero, J. C., Bourret, J. C., Sloman, K. N., Samaha, A. L., & Dallery, J. (2010). Concurrent reinforcement schedules for problem behavior and appropriate behavior: Experimental applications of the matching law. *Journal of the Experimental Analysis of Behavior, 93*, 455–469.

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

Campbell, S. B., Spieker, S., Burchinal, M., & Poe, M. D. (2006). Trajectories of aggression from toddlerhood to age 9 predict academic and social functioning through age 12. *Journal of Child Psychology and Psychiatry, 47*, 791–800.

*Carey, Y. A., & Halle, J. W. (2002). The effect of an idiosyncratic stimulus on self-injurious behavior during task demands. *Education and Treatment of Children, 25*, 131–141.

*Carter, S. L. (2010). A comparison of various forms of reinforcement with and without extinction as treatment for escape-maintained problem behavior. *Journal of Applied Behavior Analysis, 43*, 543–546. https://doi.org/10.1007/s10864-021-09453-2

Catania, A. C. (2013). *Learning* 5th (Ed.). Sloan.

Chazin, K. T., Ledford, J. R., & Pak, N. S. (2021). A systematic review of augmented input interventions and exploratory analysis of moderators. *American Journal of Speech-Language Pathology*. https://doi.org/10.1044/2020_AJSLP-20-00102

Cooper, J. O., Heron, T. E., & Heward, W. L. (2020). *Applied behavior analysis* 3rd (Ed.). Pearson.

Dart, E. H., Radley, K. C., Mason, B. A., & Allen, J. P. (2018). Addressing escape-maintained behavior for students with developmental disabilities: A systematic review of school-based interventions. *Psychology in the Schools, 55*, 295–304.

*Davis, D. H., Fredrick, L. D., Alberto, P. A., & Gama, R. (2012). Functional communication training without extinction using concurrent schedules of differing magnitudes of reinforcement in classrooms. *Journal of Positive Behavior Interventions, 14*, 162–172.

*Davis, T. N., Weston, R., Hodges, A., Uptegrove, L., Williams, K., & Schieltz, K. M. (2018). Functional communication training and demand fading using concurrent schedules of reinforcement. *Journal of Behavioral Education, 27*, 343–357.

*DeLeon, I. G., Neidert, P. L., Anders, B. M., & Rodriguez-Catter, V. (2001). Choices between positive and negative reinforcement during treatment for escape-maintained behavior. *Journal of Applied Behavior Analysis, 34*, 521–525.

Elliott, S. N. (1988). Acceptability of behavioral treatments: Review of variables that influence treatment selection. *Professional Psychology: Research and Practice, 19*, 68–80.

Gardner, A. W., Wacker, D. P., & Boelter, E. W. (2009). An evaluation of the interaction between quality of attention and negative reinforcement with children who display escape-maintained problem behavior. *Journal of Applied Behavior Analysis, 42*, 343–348.

Gast, D. L., Lloyd, B. P., & Ledford, J. R. (2018). Multiple baseline and multiple probe designs. In J. R. Ledford & D. L. Gast (Eds.), *Single case research methodology* (pp. 239–281). Routledge.

Geiger, K. B., Carr, J. E., & LeBlanc, L. A. (2010). Function-based treatments for escape-maintained problem behavior: A treatment-selection model for practicing behavior analysts. *Behavior Analysis in Practice, 3*, 22–32.

*González, M. L., Rubio, E. K., & Taylor, T. (2014). Inappropriate mealtime behavior: The effects of noncontingent access to preferred tangibles on responding in functional analyses. *Research in Developmental Disabilities, 35*, 3655–3664.

*Hoch, H., McComas, J. J., Thompson, A. L., & Paone, D. (2002). Concurrent reinforcement schedules: Behavior change and maintenance without extinction. *Journal of Applied Behavior Analysis, 35*, 155–169.

*Horner, R. H., Day, H. M., & Day, J. R. (1997). Using neutralizing routines to reduce problem behaviors. *Journal of Applied Behavior Analysis, 30*, 601–614.

Huwaldt, J. A., & Steinhorst, S. (2013). PlotDigitzer. Source code available at http://plotdigitizer.sourceforge.net/

*Ingvarsson, E. T., Hanley, G. P., & Welter, K. M. (2009). Treatment of escape-maintained behavior with positive reinforcement: The role of reinforcement contingency and density. *Education and Treatment of Children, 32*, 371–401.

Iwata, B. A. (1987). Negative reinforcement in applied behavior analysis: An emerging technology. *Journal of Applied Behavior Analysis, 20*, 361–378.

*Johnson, L., Mccomas, J., Thompson, A., & Symons, F. J. (2004). Obtained versus programmed reinforcement practical considerations in the treatment of escape-reinforced aggression. *Journal of Applied Behavior Analysis, 37*, 239–242.

*Kelley, M. E., Lerman, D. C., & Van Camp, C. M. (2002). The effects of competing reinforcement schedules on the acquisition of functional communication. *Journal of Applied Behavior Analysis, 35*, 59–63.

Kern, L., Delaney, B. A., Hilt, A., Bailin, D. E., & Elliot, C. (2002). An analysis of physical guidance as reinforcement for noncompliance. *Behavior Modification, 26*, 516–536.

Kodak, T., Miltenberger, R. G., & Romaniuk, C. (2003). A comparison of differential reinforcement and noncontingent reinforcement for the treatment of a child's multiply controlled problem behavior. *Behavioral Interventions, 18*, 267–278.

*Lalli, J. S., & Casey, S. D. (1996). Treatment of multiply controlled problem behavior. *Journal of Applied Behavior Analysis, 29*, 391–395.

*Lalli, J. S., Vollmer, T. R., Progar, P. R., Wright, C., Borrero, J., Daniel, D., … May, W. (1999). Competition between positive and negative reinforcement in the treatment of escape behavior. *Journal of Applied Behavior Analysis, 32*, 285–296.

Lane, K. L., Weisenbach, J. L., Phillips, A., & Wehby, J. H. (2007). Designing, implementing, and evaluating function-based interventions using a systematic, feasible approach. *Behavioral Disorders, 32*, 122–139.

*LaRue, R. H., Stewart, V., Piazza, C. C., Volkert, V. M., Patel, M. R., & Zeleny, J. (2011). Escape as reinforcement and escape extinction in the treatment of feeding problems. *Journal of Applied Behavior Analysis, 44*, 719–735.

Ledford, J. R., Chazin, K. T., Lane, J. D., Zimmerman, K. N., & Ayres, K. A. (2020). Single case analysis and review framework (SCARF). Retrieved from: http://ebip.vkcsites.org/scarfv2

Ledford, J. R., Lane, J. D., & Severini, K. E. (2018a). Systematic use of visual analysis for assessing outcomes in single case design studies. *Brain Impairment, 19*, 4–17.

Ledford, J. R., Lane, J. D., & Tate, R. (2018). Evaluating quality and rigor in single case research. In Jennifer R. Ledford & David L. Gast (Eds.), *Single Case Research Methodology* (pp. 365–392). Routledge.

Ledford, J. R., Whiteside, E., & Severini, K. E. (2018c). A systematic review of interventions for feeding-related behaviors for individuals with autism spectrum disorders. *Research in Autism Spectrum Disorders, 52*, 69–80.

Ledford, J. R., & Windsor, S. A. (2021). Systematic review of interventions designed to teach imitation to young children with disabilities. *Topics in Early Childhood Special Education*. https://doi.org/10.1177/02711214211007190

*Lomas, J. E., Fisher, W. W., & Kelley, M. E. (2010). The effects of variable-time delivery of food items and praise on problem behavior reinforced by escape. *Journal of Applied Behavior Analysis, 43*, 425–435.

*Lomas Mevers, J. E., Fisher, W. W., Kelley, M. E., & Fredrick, L. D. (2014). The effects of variable-time versus contingent reinforcement delivery on problem behavior maintained by escape. *Journal of Applied Behavior Analysis, 47*, 277–292.

*Long, E., Hagopian, L., Deleon, I., Marhefka, J., & Resau, D. (2005). Competing stimuli in the treatment of multiply controlled problem behavior during hygiene routines. *Research in Developmental Disabilities, 26*, 57–69.

Lugo, A. M., King, M., Lamphere, J., & McArdle, P. (2017). Developing procedures to improve therapist–child rapport in early intervention. *Behavior Analysis in Practice, 10*, 395–401.

Luiselli, J. K., Sperry, J. M., & Draper, C. (2015). Social validity assessment of physical restraint intervention by care providers of adults with intellectual and developmental disabilities. *Behavior Analysis in Practice, 8*, 170–175.

*McComas, J., Hoch, H., Paone, D., & El-Roy, D. (2000). Escape behavior during academic tasks: A preliminary analysis of idiosyncratic establishing operations. *Journal of Applied Behavior Analysis, 33*, 479–493.

Mcconnachie, G., & Carr, E. G. (1997). The effects of child behavior problems on the maintenance of intervention fidelity. *Behavior Modification, 21*, 123–158.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Prisma Group. (2010). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*. https://doi.org/10.1371/journal.pmed.1000097

*Mueller, M. M. (2002). A comparison of methods to reduce classroom behavior problems maintained by negative reinforcement (Order No. 3067241). Available from ProQuest Dissertations & Theses Global. (305569053).

*O'Reilly, M. F., Lacey, C., & Lancioni, G. E. (2000). Assessment of the influence of background noise on escapemaintained problem behavior and pain behavior in a child with Williams syndrome. *Journal of Applied Behavior Analysis, 33*, 511–514.

*O'Reilly, M., Sigafoos, J., Lancioni, G., Edrisinha, C., & Andrews, A. (2005). An examination of the effects of a classroom activity schedule on levels of self-injury and engagement for a child with severe autism. *Journal of Autism and Developmental Disorders, 35*, 305–311.

Owen, T. M., Luczynski, K. C., Rodriguez, N. M., & Fisher, W. W. (2021). Assessment of children's and caregivers' preferences for treatments for escape-maintained problem behavior. *Journal of Applied Behavior Analysis*. https://doi.org/10.1002/jaba.817

*Pace, G. M., Ivancic, M. T., & Jefferson, G. (1994). Stimulus fading as treatment for obscenity in a brain-injured adult. *Journal of Applied Behavior Analysis, 27*, 301–305.

Park, K. L., & Scott, T. M. (2009). Antecedent-based interventions for young children at risk for emotional and behavioral disorders. *Behavioral Disorders, 34*, 196–211.

*Phillips, L. A., Fritz, J. N., Rettig, L. A., & Martin, K. E. (2019). Diaphragmatic breathing as treatment for escapemaintained aggression. *Behavioral Interventions, 34*, 216–230.

*Piazza, C. C., Fisher, W. W., Hanley, G. P., Remick, M. L., Contrucci, S. A., & Aitken, T. L. (1997). The use of positive and negative reinforcement in the treatment of escape-maintained destructive behavior. *Journal of Applied Behavior Analysis, 30*, 279–298.

*Piazza, C. C., Hanley, G. P., Fisher, W. W., Ruyter, J. M., & Gulotta, C. S. (1998). On the establishing and reinforcing effects of termination of demands for destructive behavior maintained by positive and negative reinforcement. *Research in Developmental Disabilities, 19*, 395–407.

Pustejovsky, J. E. (2019). clubSandwich: Cluster-robust (Sandwich) variance estimators with small-sample corrections. R package version 0.3.4. https://cran.rproject.org/web/packages/clubSandwich/index.html

Pustejovsky, J. E. (2018). Using response ratios for meta-analyzing single-case designs with behavioral outcomes. *Journal of School Psychology, 68*, 99–112.

R Core Team. (2016). R: A language and environment for statistical computing. Vienna, Austria: R foundation for statistical computing. http://www.R-project.org/

Rajaraman, A., Hanley, G. P., Gover, H. C., Staubitz, J. L., Staubitz, J. E., Simcoe, K. M., & Metras, R. (2021). Minimizing escalation by treating dangerous problem behavior within an enhanced choice model. *Behavior Analysis in Practice,* 1–24.

*Robertson, R. E., Wehby, J. H., & King, S. M. (2013). Increased parent reinforcement of spontaneous requests in children with autism spectrum disorder: Effects on problem behavior. *Research in Developmental Disabilities, 34*, 1069–1082.

*Romaniuk, C., Miltenberger, R., Conyers, C., Jenner, N., Jurgens, M., & Ringenberg, C. (2002). The influence of activity choice on problem behaviors maintained by escape versus attention. *Journal of Applied Behavior Analysis, 35*, 349–362.

*Schieltz, K. M., Wacker, D. P., & Romani, P. W. (2016). Effects of signaled positive reinforcement on problem behavior maintained by negative reinforcement. *Journal of Behavioral Education, 26*, 137–150.

Schmidt, J. D., Luiselli, J. K., Rue, H., & Whalley, K. (2013). Graduated exposure and positive reinforcement to overcome setting and activity avoidance in an adolescent with autism. *Behavior Modification, 37*, 128–142.

*Schmidt, J. D., Shanholtzer, A., Mezhoudi, N., Scherbak, B., & Kahng, S. (2014). The utility of a brief experimental analysis for problem behavior maintained by escape from demands. *Education and Treatment of Children, 37*, 229–247.

Shadish, W. R., Zelinsky, N. A., Vevea, J. L., & Kratochwill, T. R. (2016). A survey of publication practices of single-case design researchers when treatments have small or large effects. *Journal of Applied Behavior Analysis, 49*, 656–673.

*Shirley, M. J., Iwata, B. A., Kahng, S. W., Mazaleski, J. L., & Lerman, D. C. (1997). Does functional communication training compete with ongoing contingencies of reinforcement? An analysis during response acquisition and maintenance. *Journal of Applied Behavior Analysis, 30*, 93–104.

*Slocum, S. K., & Vollmer, T. R. (2015). A comparison of positive and negative reinforcement for compliance to treat problem behavior maintained by escape. *Journal of Applied Behavior Analysis, 48*, 563–574.

Stormont, M., Beckner, R., Mitchell, B., & Richter, M. (2005). Supporting successful transition to kindergarten: General challenges and specific implications for students with problem behavior. *Psychology in the Schools, 42*, 765–778.

*Stuesser, H. A., & Roscoe, E. M. (2020). An evaluation of differential reinforcement with stimulus fading as an intervention for medical compliance. *Journal of Applied Behavior Analysis, 53*, 1606–1621.

Trump, C. E., Ayres, K. M., Quinland, K. K., & Zabala, K. A. (2019). *Differential reinforcement without extinction: A review of the literature*. Research and Practice. Advance online publication.

United Kingdom Department of Health. (2015). Positive and proactive care: reducing the need for restrictive interventions. Retrieved from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/300293/JRA_DoH_Guidance_on_RP_web_accessible.pdf

*Vasquez, S., Brewer, A., Leon, Y., & Vasquez, J. (2017). The effects of advance notice on problem behavior occasioned by interruptions of an ongoing activity in a young girl with autism. *Behavior Analysis in Practice, 10*, 417–421.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*, 1–48.

*Wilder, D., Normand, M., & Atwell, J. (2005). Noncontingent reinforcement as treatment for food refusal and associated self-injury. *Journal of Applied Behavior Analysis, 38*, 549–553.

Wolf, M. M. (1978). Social validity: The case for subjective measurement or how applied behavior analysis is finding its heart. *Journal of Applied Behavior Analysis, 11*, 203–214.

*Zarcone, J. R., Iwata, B. A., Mazaleski, J. L., & Smith, R. G. (1994a). Momentum and extinction effects on self-injurious escape behavior and noncompliance. *Journal of Applied Behavior Analysis, 27*, 649–658.

*Zarcone, J. R., Iwata, B. A., Smith, R. G., Mazaleski, J. L., & Lerman, D. C. (1994b). Reemergence and extinction of self-injurious escape behavior during stimulus (instructional) fading. *Journal of Applied Behavior Analysis, 27*, 307–316.