Systematic Review of Interruption and Redirection Procedures for Autistic Individuals

Jennifer R. Ledford (Jennifer.r.ledford@vanderbilt.edu)

Vanderbilt University

Cary Trump (cary.trump@unco.edu)

University of Northern Colorado

Kate T. Chazin (kate.chazin@vanderbilt.edu)

Sienna A. Windsor (sienna.windsor@vanderbilt.edu)

Brittany P. Bennett (brittany.p.bennett@vanderbilt.edu)

Vanderbilt University

& Kara Wunderlich (kwunderlich@rollins.edu)

Rollins College

Abstract

Interruption and redirection (IR) procedures involve systematically disrupting an undesirable behavior and prompting an individual to engage in an alternative behavior (e.g., Ahearn et al., 2007). These procedures have been frequently assessed for reducing stereotypy for autistic individuals. The purposes of this review were to determine (a) for whom and under what conditions physical and verbal IR procedures are assessed; (b) what outcomes are measured; (c) whether results vary by intervention characteristics, dependent variable types, design quality, and publication status; and (d) the extent to which generalization and maintenance outcomes are measured and show positive effects. IR interventions were moderately successful for reducing stereotypy but were less effective for improving appropriate vocalizations and functional engagement. Generalization and maintenance outcomes, when measured, were poor. Results indicate the need for alternative or augmentative procedures that focus on engagement in meaningful interactions and activities.

Systematic Review of Interruption and Redirection Procedures

Autistic individuals¹ often engage in repetitive behaviors that are socially atypical. These may be repeated or unusual vocal or motor behaviors (generally referred to as stereotypy), repetitive actions with objects that are more complex (e.g., hoarding, ordering, or arranging specific items), or self-injurious behaviors (e.g., pica, head banging). Stereotypy is often automatically maintained (Beavers et al., 2013), although the definition of stereotypy typically focuses on form rather than function (Hagopian & Toole, 2009). Some autistic self-advocates have identified at least a subset of these behaviors to be advantageous because they serve a self-regulation function (Joyce et al., 2017; Manor-Binyamini & Schreiber-Divon, 2019) and emerging research supports this possibility (Heathers et al., 2019). Moreover, at least one study has found that vocal stereotypy may have facilitative effects on learning (Sherer & Schreibman, 2005).

Despite these conceivable benefits, repetitive behaviors may also have the potential to have negative impacts on learning or social outcomes (Morrison & Rosales-Ruiz, 1997) or may be physically dangerous. Causal paths are unclear, but people who engage in stereotypy tend to have more pronounced autistic symptomology (Powell et al., 2017), lower IQs (Goldman et al., 2008), and perform lower on measures of social and adaptive skills (Matson et al., 1997). Given these characteristics, the population of autistic people who engage in persistent stereotypy may be among the most vulnerable. Because these individuals may be less likely to effectively self-advocate, it is especially important to identify whether typical procedures used in their educational and behavioral treatments are effective and acceptable.

---

¹ We use identity-first language purposefully, given many self-identified autistic people prefer this (e.g., Bury et al., 2020; Kenny et al., 2016) and given guidance on using preferred language in APA 7. We do this despite academic tradition for using person-first language, and while acknowledging that some individuals may prefer person-first language.

Treating Stereotypy

When considerable negative impacts are likely (e.g., harm to self or others), the benefits of treating repetitive behaviors (e.g., promoting safety) may outweigh the loss of benefits those behaviors could provide. Selecting and implementing interventions for automatically maintained behaviors may be more complex than for socially mediated ones, although researchers and practitioners have demonstrated that some people exhibit less stereotypy when competing stimuli are available and when differential reinforcement procedures are used (e.g., differential reinforcement of alternative behavior; Akers et al., 2020). Unfortunately, as with other interventions identified as evidence-based (cf. Ledford et al., 2021), there are also a multitude of instances in which these procedures do not lead to improvements in stereotypy.

When selecting treatments that target repetitive behaviors, consideration of replacement behaviors is important. One review suggested that when stereotypy decreases in intervention conditions, reallocation towards other behaviors occurs—to both socially desirable and undesirable behaviors (Lanovaz et al., 2013). Thus, it is critical to consider response strengthening when treating repetitive behaviors (Cunningham & Schreibman, 2008). That is, when interventionists determine a need to intervene on stereotypical behaviors, they should focus on increasing functional engagement and appropriate social interactions in addition to decreasing repetitive behaviors that are likely to cause harm or disrupt learning.

In addition to complexities associated with selecting an effective intervention, treatment of repetitive behaviors requires consideration of ethical and social validity issues. For example, additional work is needed to determine under what conditions intervening on stereotypy is appropriate and acceptable to all stakeholders, including the individual engaging in stereotypy. Thus, treating stereotypy for autistic individuals is complex and requires considerable expertise

and deliberation regarding the extent to which the treatment is appropriate, acceptable, and effective at reducing problematic behaviors while improving desirable ones.

Interruption and Redirection Interventions

One intervention that has been identified by some authors as evidence-based for treating repetitive behaviors includes interrupting the behavior and then redirecting the individual to engaging in a different behavior (interruption and redirection; IR). This can take the form of physically interrupting the behavior (response blocking) or vocally interrupting the behavior (e.g., name call) and then physically redirecting the individual to do something else (e.g., manual guidance to play with available toys) or vocally redirecting the individual to do something else (e.g., telling the child to imitate an action or engage in a vocal response). Generally, vocal interruption is followed by vocal redirection (VIR; often called response interruption and redirection, RIRD) and physical interruption is followed by physical redirection (PIR, sometimes called RIRD). VIR sometimes includes subsequent physical guidance (i.e., vocal interruption, followed by least-to-most prompting to engage in the redirected behavior; VIR+P), blurring the lines between PIR and VIR. We will refer to the IR procedures as VIR, VIR+P, or PIR because these names are more descriptive and precise than the common term RIRD that often, but not always, refers to VIR with or without physical guidance.

Previous Reviews

A number of researchers have reviewed IR interventions, although their purposes and the included population of studies have varied somewhat. First, a review by Lydon and colleagues (2013) separated RIRD from redirection alone, with RIRD corresponding with the procedure described as VIR above. Because redirection generally requires that a behavior is interrupted (e.g., interaction occurs contingent on the behavior and requires the person to engage in a

different competing alternative behavior), the differences between what they referred to as RIRD and redirection alone may primarily be topography of the IR (e.g., VIR versus PIR). This review included 18 peer-reviewed sources and categorized IR interventions as *not evidence based* for decreasing challenging behaviors (including repetitive behaviors). Effectiveness for other behaviors was not assessed.

A more recent review (Steinbrenner et al., 2020) defined RIRD as introducing a *distractor* to reduce an *interfering* behavior and included 29 peer-reviewed sources that used VIR, PIR, and response blocking alone (i.e., no redirection occurred). They identified these procedures as being *evidence based* for improving a wide variety of outcomes: academic behaviors, adaptive behavior, challenging behaviors, communication, motor skills, play skills, school readiness (e.g., engagement), and social behaviors. However, the procedures and resulting findings from this large-scale review have been identified as potentially problematic (Leaf, 2021; Ledford et al., 2021). Importantly, because this review only included studies with positive outcomes, it does not provide information about negative or equivocal findings.

Another recent review (Dowdy et al., 2020) searched for studies that used terms associated with RIRD and included those published in peer-reviewed journals ($n = 25$) and unpublished (e.g., dissertations and theses; $n = 13$) sources. They found RIRD was effective for reducing automatically-maintained behavior and that published studies had similar quality but slightly better outcomes than unpublished studies (indicating potential publication bias). They did not evaluate other outcomes (e.g., engagement, communication). Given differences in search strategies, methods, included studies, and terminology, an updated and comprehensive review is needed, including a variety of dependent variables, published and unpublished sources, and including all studies that incorporate IR procedures (e.g., VIR, VIR+P and PIR).

Contributions of this Review

The purpose of this review is to describe the current state of research supporting IR intervention use for autistic people. This review will advance current understanding of IR interventions because it (a) includes both published and unpublished sources, (b) includes only designs that offer convincing assessments (e.g., three potential demonstrations of effect), (c) assesses outcomes in relation to quality/rigor, and (d) includes interventions with interruption and redirection components, separated into topographical categories. Research questions guiding this review were: (1) For whom and under what conditions are IR interventions assessed? (2) Do outcomes vary by topography type (PIR, VIR, VIR+P)? (3) What dependent variables are measured, and do outcomes vary by dependent variable type? (4) What is the overall rigor/quality of designs assessing IR interventions and do outcomes vary by rigor/quality or publication status? (5) How robust are generalization and maintenance outcomes? It is important to note that questions regarding social and ecological validity of IR interventions are also important; these questions are explored in a separate manuscript.

Method

We screened and coded single case design studies of IR interventions for autistic individuals. We considered *sources* as individual articles or dissertations written by one or more authors that include one or more assessments of IR interventions. A source may have included only one *design* (e.g., a source that includes a single A-B-A-B assessment for one dependent variable for one participant) or multiple *designs* (e.g., a source that includes an A-B-A-B comparison for two dependent variables for one participant or an A-B-A-B comparison for one dependent variable for two participants would each have two designs). Agreement data (89-94% across steps) for screening and coding is available in Supplemental Materials.

Inclusion Criteria

We included designs that compared IR interventions to baseline or business-as-usual conditions for treatment of challenging behavior, stereotypy, or a category of repetitive behaviors we described as hoarding, ordering, and arranging (HOA); see definitions below. We included sources that contained one or more designs with adequate rigor to confidently interpret effects. These designs included (a) three potential demonstrations of effect (Ledford & Gast, 2018; WWC, 2020) and (b) at least three data points per condition *or* author rationale for fewer measurement occasions based on dangerous behavior. Within a source, we coded designs that were eligible, and did not code ineligible designs (e.g., given two A-B-A-B designs in a source, one for an autistic child and one with a child who did not have autism, we coded only the first design). For multiple baseline designs across participants, we included designs with at least two autistic participants. We excluded designs that were conducted for assessment rather than intervention (e.g., using eligible intervention components during a functional analysis).

Search and Screening

We used a two-pronged search strategy to search for eligible sources. First, we searched PsycINFO and ProQuest Dissertations and Theses on September 7, 2020 (see search terms and PRISMA diagram in Supplemental Materials). We exported potential sources ($n = 399$) to Raayan (Ouzzani et al., 2016) for initial screening using titles and abstracts with two independent coders (100% double screened). Sources identified as potentially eligible (i.e., “screened in”) by either coder were reviewed at the full text level ($n = 126$). Following completion of the electronic search, we conducted a backwards search for four previous reviews with similar but distinct purposes (Dowdy et al., 2020; Lydon et al., 2013; Spencer & Alkhanji, 2018; Steinbrenner et al., 2020). Two independent reviewers screened all articles at the title/abstract level except those that

were duplicates of sources found in the database search. If a source was identified as potentially eligible by at least one coder in the backwards screening, it was included in the full text review ($n = 25$). Two independent coders then reviewed full text sources ($n = 151$; see PRISMA diagram).

Descriptive Coding

We coded descriptive information to characterize contexts and components of IR interventions. We coded some information by *source* (e.g., journal name, year of publication), and other information by *participant* (e.g., participant pseudonym and age), but most information was coded for each included *design* (e.g., intervention type, dependent variable). The coding document available [here](#) includes definitions and data for each code.

General Information

For each *source*, we coded the year of publication, whether data were published in a peer-reviewed journal or were unpublished (dissertations and theses), and the journal in which it was published (if applicable). Prior to other descriptive coding, we conducted a separate step (with 100% double coding) to identify each eligible design in every source. For each *design*, we coded the design type (e.g., A-B-A-B, alternating treatments, multiple baseline across participants), the included participants (e.g., pseudonyms) in that design, dependent variable type, and other information that would help to differentiate designs (e.g., if two different A-B-A-B designs were conducted in separate settings, we identified the settings in this column to facilitate coding and agreement assessment).

Dependent Variables

For each eligible *design* in a source, we coded the type of dependent variable measured as vocal stereotypy, motor stereotypy, HOA (hoarding, ordering, and arranging), dangerous

behaviors (e.g., pica), other challenging behavior, functional engagement, appropriate vocalizations, and task completion. An *other* category was also coded (e.g., emotional responses, item contact). Some designs included measurement of a single dependent variable comprising multiple categories (e.g., motor plus vocal stereotypy); in these cases, we coded all dependent variables that applied. For analysis, we combined motor stereotypy and HOA codes, given HOA was targeted relatively rarely ($n = 2$; in both instances, definitions included both motor stereotypy and HOA behaviors).

We also recorded whether data were measured as count, momentary time sampling or partial interval recording, percentage of opportunities, or duration. In addition, we recorded whether authors followed graphing standards by including full range of potential values, when a percentage was reported (i.e., vertical axis represents 0 to 100%). Finally, we coded whether measurement was continuous or interrupted. *Continuous* data collection referred to collecting data on dependent variables throughout all sessions. *Interrupted* data collection referred to collecting data only during the portions of the session when treatment was not being actively implemented (e.g., a stopwatch was stopped when interruption commenced and restarted when the redirection was complete). That is, stereotypy that occurred while the implementer was interrupting and redirecting was not measured.

Participant and Context

For each eligible *participant* in a source, we coded the identifier (e.g., pseudonym), age, gender, disability (autism, to confirm eligibility; intellectual disability and vision impairment, as both are also associated with stereotypy, and other disabilities identified). We recorded implementer as researcher, parent, teacher or other practitioner who is endogenous (i.e., serves the child outside of the study context), not reported, and other. We coded ambiguous terms as

“not reported”—for example, the term “therapist” could refer to someone who implements programming outside of the research study or as someone who is implementing a research protocol. For each participant, we coded the intervention setting (e.g., separate school or residential facility, clinical facility or hospital inpatient/outpatient, home, neighborhood school).

Intervention

For each eligible *design* in a source, we coded information about baseline conditions: this included what materials were available (none, preferred leisure, moderately preferred leisure, work tasks, low or non-preferred items, leisure items with unclear preference, materials specifically associated with low or high levels of stereotypy, not reported, or other). We also coded adult interactions contingent on stereotypy and *not* contingent on stereotypy. For the former, we coded ignoring, reprimands, removal of materials, no adult present, not reported, or other. For the latter, we coded play (responsive interactions and engagement), respond only (appropriate or minimal responses to child initiations), prompts to work or play, ignore, no adult present, not reported, and other. We also coded baseline session time when reported.

For intervention conditions, we coded the type of intervention (PIR, VIR, VIR+P, other/mixed), whether preferred materials were removed contingent on targeted behaviors, and whether intervention conditions were signaled. Definitions of PIR, VIR, VIR+P are in the introduction section and definitions and examples are available in Supplemental Materials. Mixed and other interventions were coded when the interruption and redirection components were different topographies (e.g., vocal interruption, physical redirection) or multiple topographies (e.g., physical and vocal interruption). VIR+P was different from mixed in that VIR+P included vocal interruptions and redirections and physical topographies were only

introduced as error correction, when needed (i.e., the child did not necessarily contact physical guidance). Mixed interventions included various topographies regardless of errors.

For intervention conditions, we also coded the topography of behavior for which IR interventions were initiated and the expected topography of the response to redirection and we determined whether these were matched (e.g., vocal stereotypy was interrupted, and the expected response was vocal), unmatched (e.g., vocal stereotypy was interrupted and the expected response was motor), or partially matched (e.g., vocal and motor stereotypy were interrupted and the expected response was vocal). In addition, we coded whether the stop criterion was a certain number of compliances (including prompted compliances) or commands (even if not complied with). We also recorded whether redirections were contextually relevant (i.e., related to the task at hand) or not contextually relevant (i.e., unrelated). We also coded if the type of redirection was unclear or whether it was not possible for the redirection to be contextually relevant because the child did not have access to people or materials to engage with (e.g., no materials, an adult who ignored non-target behaviors).

SCARF Coding

We used the Single Case Analysis and Review Framework (SCARF; Ledford et al., 2020) to evaluate quality and rigor of included designs and to assess outcomes. Information about SCARF coding and interpretation is available [here](#) (Ledford et al., 2020).

Quality and Rigor

The SCARF evaluates internal and external validity across categories, including the extent to which authors provide sufficient descriptions of participants, conditions, and dependent variables; whether reliability data are sufficiently collected, reported, and indicative of reliable implementation and data collection (i.e., fidelity and interobserver agreement); whether there is

evidence of social and ecological validity; whether generalization and maintenance data are collected; and whether sufficient data exist. This final category requires absence of problematic patterns which could limit conclusions drawn from the data; thus, a study could be eligible with three data points in each condition but score poorly in this category (e.g., when baseline data have a therapeutic trend). The weighted average across categories (with internal validity items rated more heavily than external validity items) ranges from 0–4.

Outcomes

The SCARF also allows for coding of three types of outcomes, all based on visual analysis: primary outcomes, generalization, and maintenance. Visual analysis is used to evaluate the extent to which changes are consistent and unequivocal on a five-point scale (0-4). Scores of 0 or 1 are indicative of non-effects and no functional relations, a score of 2 is indicative of mixed effects, and scores of 3 or 4 are indicative of effects that are consistent with a functional relation determination, with scores of 3 indicating potentially weak effects and scores of 4 indicating robust positive effects (e.g., minimal or no overlap between conditions, immediate changes).

Maintenance and generalization outcomes are interpreted similarly to primary outcomes, with scores of 3 and 4 indicating unequivocal positive effects. For both generalization and maintenance, we coded whether these conditions were different from intervention (either in context or time) but in which *intervention was still implemented*, or whether *intervention was withdrawn*. For example, some researchers measured generalization to a new environment or implementer, but intervention procedures were implemented in these different contexts (generalization, treatment in place). Others measured generalization to new contexts, under baseline or business as usual conditions (generalization, treatment withdrawn). Similarly for maintenance, some researchers measured behavior following withdrawal of intervention

procedures (maintenance, treatment withdrawn) while others measured whether behavior maintained after several weeks of continued intervention (maintenance, treatment in place). When no intervention was in effect, we coded whether the maintenance was *within session* or *across sessions*; that is, whether maintenance of effects occurred short term (e.g., 5 min after the IR intervention was discontinued) or long term (e.g., days or weeks after the IR intervention was discontinued).

SCARF outcomes (primary, generalization, and maintenance) are depicted via scatterplot. For primary outcomes, quality and rigor scores are depicted on the vertical axis while outcome scores are depicted on the horizontal axis. Thus, high-quality studies with positive outcomes are in the top right of the scatterplot, while high-quality studies with mixed or poor outcomes are in the bottom right. Studies with low quality are depicted on the left side of the graph. There are no published guidelines for determining a quality “cut-off” score, but generally, studies that are further right on the graph should be considered to provide more believable evidence, regardless of outcomes (e.g., believable evidence of effects or non-effects). Other studies have used 1.7, 1.8, or 2.0 as quality cut-offs (Chazin et al., 2021; Gibbs et al., 2021; Ledford & Windsor, 2021). Given our requirements for data and potential demonstrations, we elected to include data from all studies in our analyses, and to assess whether evidence differed between lower and higher quality studies.

Results

General Information

A total of 53 sources were identified ($n = 33$ published, 20 unpublished; between search and completion of this manuscript, 1 unpublished dissertation was published—we left this source categorized as *unpublished* given it was identified as such in the original search). Across sources,

there were 291 designs and 133 autistic participants. A total of 45 sources were identified in the electronic review, and 7 novel sources were identified via backwards searches. One source, published after the search was conducted, was identified when a journal table of contents was automatically sent via email to the first author (see PRISMA diagram in Supplemental Materials). Agreement data for descriptive coding are available in Supplemental Materials (average of 94-100% across domains) and SCARF coding (average of 92-95% across domains).

Descriptive Information

General

The most used designs were A-B-A-B ($n = 230$) and ATD ($n = 53$), with remaining designs including multiple baseline or multiple probe designs across participants ($n = 8$). Papers were published beginning in 1996, with multiple studies published (journals) or completed (theses/dissertations) in each year from 2009-2020. Studies were published in the *Journal of Applied Behavior Analysis* ($n = 14$; including one that was originally unpublished) or *Behavioral Interventions* ($n = 11$) more than half the time, with remaining articles ($n = 9$) published in 8 different journals. Most designs assessed VIR ($n = 151$) or VIR+P interventions ($n = 105$), with fewer assessing PIR ($n = 13$) or mixed interventions ($n = 22$).

Dependent Variables and Measurement Systems

The most common dependent variable measured was vocal stereotypy ($n = 134$), followed by motor stereotypy ($n = 50$), functional engagement ($n = 33$), appropriate vocalizations ($n = 44$), combined vocal and motor stereotypy ($n = 13$), dangerous behaviors ($n = 7$), task completion ($n = 4$), and other behaviors ($n = 6$). Vocal stereotypy, motor stereotypy, functional engagement, and dangerous behaviors were measured equally during PIR interventions ($n = 3$ for each). Vocal stereotypy ($n = 66$), appropriate vocalizations ($n = 38$), and

motor stereotypy ($n = 24$) were measured most often for VIR interventions. Vocal stereotypy ($n = 65$), motor stereotypy ($n = 24$), and functional engagement ($n = 14$) were measured most often in VIR+P interventions. Mixed IR interventions primarily measured motor stereotypy ($n = 9$) and functional engagement ($n = 5$). See Table 1 for more information.

In terms of measurement procedures, continuous measurement was used in 138 designs, while interrupted measurement was used in 127 designs. Procedures were not sufficiently described for categorization in 22 designs, and in 4 designs, data were presented for treatment time only. Most designs for PIR ($n = 11$, 85%) and mixed IR interventions ($n = 15$, 68%) used continuous measurement while fewer VIR ($n = 59$, 39%) and VIR+P ($n = 42$, 40%) designs included continuous measurement.

Participants

Per inclusion criteria, all participants met criteria for an autism spectrum disorder. Most were boys ($n = 108$, 80%) and had no reported secondary disabilities ($n = 119$, 89%). Few were reported to have intellectual disability ($n = 13$), ADHD ($n = 2$, other category), or other disabilities ($n = 2$; some participants identified with multiple additional disabilities). Participants' ages ranged from 2 to 36, with a mean age of 10 years and a median age of 9 years. Two participants were older than 21 years, 40 were secondary school-aged (12-21 years), 68 were primary school-aged (6-11 years), and 23 were young children (5 years or younger).

Contexts

For nearly half of participants ($n = 63$), the role of the implementer was unclear. For 68 participants for whom the role was clear, most were researchers ($n = 49$, 72%), with some endogenous practitioners such as teachers or clinicians who served the child outside the context of the study ($n = 18$, 26%), and one parent. Nearly half ($n = 59$, 44%) of individuals participated

in studies in educational settings that were designed to provide services only to individuals with disabilities (specialized schools, treatment rooms in residential centers, etc.). Fewer participated in clinic settings ($n = 22$), their homes ($n = 14$), self-contained classrooms in local schools that enrolled students without disabilities ($n = 10$), or hospital settings ($n = 7$). For 20 participants, settings were not sufficiently described for categorization. No investigations took place in inclusive, general education, or community settings.

Most dependent variables were measured via duration recording and reported as a percentage of session or number of seconds ($n = 158$). Some were measured as a percentage of intervals, using partial interval recording ($n = 67$) or measured as count and reported as number or rate ($n = 50$). Eight were measured using whole interval recording and five were measured using momentary time sampling. For three, the system reported via text did not correspond with reported units on the graph. A non-trivial minority of graphs depicting percentage did not depict the entire range (e.g, included a scale of 0 to 30% on the x-axis), potentially distorting visual analysis and resulting in overestimates of behavioral changes ($n = 38$ of 230 designs reporting percentages, 17%; Ledford & Gast, 2018). Measurement systems varied based on outcome; the majority of vocal stereotypy (74%), motor stereotypy (56%), and combined vocal/motor stereotypy (62%) were measured via duration recording. The majority of appropriate vocalizations (70%) and dangerous behaviors (71%) were measured via count. No single type was majority for functional engagement, with duration (45%) and partial interval recording (24%) used most often (see Supplemental Materials).

Baseline Information

It is important to understand baseline conditions when interpreting intervention outcomes. Baseline contexts were generally unlikely to be therapeutic, but procedures varied

across studies. The most common types of material available were leisure items ($n = 169$ designs) or no materials ($n = 51$ designs), with fewer including work materials ($n = 29$ designs) or not describing materials ($n = 38$ designs). For the remaining designs ($n = 4$), authors specified materials were associated with high stereotypy, that materials varied, or that typical classroom materials were used.

In 112 designs, implementers ignored participants during baseline conditions. In 85 designs, implementers ignored participants except for minimally responding to participant initiations. In 22 designs, implementers praised engagement and responded to participant initiations. In 28 designs, adults provided directives or prompts, sometimes also providing praise for compliance. In 20 designs, the type of implementer interaction available was not specified, and in one design, no implementer was present. We classified 23 designs as “other,” which included reinforcing mands, vocalizations, or correct responding.

Intervention Characteristics

We categorized interventions as PIR, VIR, VIR+P, and mixed IR interventions. PIR included physical interruption of motor behaviors (i.e., blocking) followed by physical prompts to engage in a different behavior ($n = 13$). VIR included vocal interruption of vocal or motor stereotypy (e.g. name call) followed by vocal redirection (e.g., ask a question, provide an instruction to complete a simple motor task; $n = 151$), while VIR+P included vocal interruption and vocal redirection that was followed by physical prompts for error correction (when needed; $n = 105$). Mixed IR interventions did not fall into these categories and included interventions in which vocal interruptions were followed by physical redirections, physical interruptions were followed by vocal redirections, or vocal and physical interruptions and/or redirections were used concurrently ($n = 22$).

Topography: Target behaviors and redirections. PIR was used primarily to interrupt motor stereotypy ($n = 10$, 77%), while VIR and VIR+P were used primarily to interrupt vocal stereotypy ($n = 111$, 74% and $n = 70$, 67%, respectively; see Supplemental Materials). When using PIR and VIR+P, implementers primarily redirected children to engage in motor behaviors ($n = 11$, 85% for PIR; $n = 70$, 67% for VIR+P), while they redirected to vocal ($n = 84$, 56%) behaviors for VIR. The topography of redirection matched the target behavior topography for 182 designs; it was unmatched for 53 designs and partially matched for some ($n = 17$). It was insufficiently described for categorization for 39 designs.

Contextual relevance. The use of contextually relevant redirections occurred for only 55 designs ($n = 19\%$). The proportion of designs including contextually relevant redirections varied considerably across intervention types; PIR included mostly contextually relevant redirections ($n = 12$, 92%) while both VIR and VIR+P rarely included redirections that were contextually relevant ($n = 15$, 10% and 13 , 12%, respectively). Mixed IR interventions included 15 designs with contextually relevant redirections (68%; see Table 2).

Signals and stop criteria. We coded whether implementers signaled to participants when intervention conditions were in effect, when compared to baseline conditions. Overall, 120 designs included signaled intervention conditions ($n = 5$, 38% for PIR interventions; $n = 58$, 38% for VIR interventions; $n = 50$, 48% for VIR+P interventions; and $n = 7$, 32% for mixed IR interventions). For PIR interventions, no stop criteria were reported for any study. For VIR, the most common stop criterion was presentation of a certain number of demands (usually three; $n = 112$, 74%) and VIR+P, the most common stop criterion was compliance with a certain number of commands (usually three; $n = 99$, 94%; see Table 3).

Additional components. We coded whether authors provided contextually appropriate responses (e.g., answering questions or providing praise) or other planned consequences (e.g., systematically planned putative reinforcers) for social initiations, complying with redirections, or other desirable participant behaviors (e.g., engagement). In a little more than half of the designs ($n = 155$), implementers provided consequences for participant initiations, almost always ($n = 145$) by providing contextually appropriate responses such as praise. For most designs ($n = 173$), implementers provided praise or acknowledgement for compliance with redirection, while 11 provided tokens or edibles in addition to praise; in remaining designs, implementers either specified no consequences were provided or did not specify consequences. Implementers in 69 designs provided consequences for other behaviors, most often engagement behaviors, and most often via contextually appropriate responding (e.g., praise). Some designs ($n = 57$) included other components, such as DRO ($n = 7$), prompts for engagement or other alternative behaviors ($n = 32$), prompts to mand for preferred items ($n = 16$), or visual supports for task completion ($n = 2$). In some designs, a negative punishment component was included, with the removal of preferred items occurring in 54 designs (19%), most commonly in VIR ($n = 23$) and VIR+P interventions ($n = 26$).

SCARF Coding

Quality and Rigor

Quality and rigor scores ranged from 0.97 to 3.36. Average scores varied slightly by type, with PIR studies average a score of 2.24 (range: 0.97-3.06), VIR studies averaging 2.23 (range: 1.56-3.36), VIR+P studies averaging 2.17 (range: 1.41-2.75), and mixed IR studies averaging 2.29 (range: 1.61-2.81). On average, the lowest-scoring category related to internal validity was procedural fidelity data (i.e., collecting and reporting adequate data on procedural reliability) and

the lowest-scoring categories related to external validity were generalization and maintenance, measured in few studies. See SCARF coding available [here](#).

Primary Outcomes

Overall effects, quality, and publication status. The most common outcome score across interventions and dependent variables was a 4 (68% of designs), consistent with strong functional relations. As shown in Supplemental Materials, non-effects, inconsistent effects, and positive effects were present for both low-rigor (e.g., outcomes depicted farther to the left of the scatterplot) and high-rigor (e.g., outcomes depicted farther to the right of the scatterplot) designs. Score dispersion was similar for published and unpublished studies, with 68% of designs in published sources and 67% of designs in unpublished sources supporting a functional relation determination (SCARF scores of 3 or 4). See Supplemental Materials. There was a difference in functional relation determination when quality was considered, with 60% of designs in the lowest quartile (i.e., the 72 designs with the lowest quality and rigor scores) supporting a functional relation determination and 75% of designs in the highest quartile supporting a functional relation determination. Similarly, in the lowest quartile, 39% of designs received a score of 0 or 1 while in the highest quartile, only 22% of designs received those scores (see Table 4).

Effects by intervention type and dependent variable. The percentage of designs suggestive of functional relations was similar for PIR (69%), VIR (68%), and VIR+P (65%). It was slightly higher for mixed IR interventions (82%), although it is difficult to interpret this value because this category represented heterogeneous procedures across studies (see Table 5 and Figures 1-5). The effectiveness varied considerably by dependent variable (see Table 6), with IR procedures being highly effective for reducing motor stereotypy (functional relations in 90% of

designs), fairly effective for reducing vocal stereotypy (75%), and less effective for increasing functional engagement (55%) or appropriate vocalizations (28%). Across intervention types and dependent variables, some patterns emerged, with PIR being less effective than VIR and VIR+P at reducing stereotypy, although this conclusion is tempered because there were few PIR studies. In addition, VIR and VIR+P interventions were particularly ineffective at increasing functional engagement and appropriate vocalizations. For dependent variables related to vocalizations, VIR+P was less effective than VIR—perhaps because VIR+P interventions more often used motor redirection responses.

Effects by intervention, measurement type, and signals. For VIR and VIR+P interventions, we analyzed whether differences in functional relation determination existed based on whether intermittent or continuous measurement was used. There were an insufficient number of studies in each category to do this analysis for PIR and mixed IR interventions. For vocal stereotypy, motor stereotypy, and appropriate vocalizations, using interrupted measurement resulted in a higher percentage of functional relations. For functional engagement, there was a greater percentage of functional relations identified when continuous measurement was used (see Supplemental Materials). As shown in Supplemental Materials, signaled interventions resulted in more functional relations than unsignaled interventions for appropriate vocalizations (50% versus 18%) and motor stereotypy (100% versus 83%) while unsignaled interventions resulted in more positive outcomes for appropriate vocalizations (91% versus 60%).

Generalization and Maintenance Outcomes

Only 36 designs (12%) included measurement of generalization. Sixteen received scores of 1 for generalization (consistent non-effects or counter-therapeutic effects), five received scores of 2 (at least one effect and at least one non-effect), one received a score of 3 (weak

positive effects), and 14 received scores of 4 (strong positive effects). Only fourteen designs measured generalization *in the absence of treatment*. That is, generalization was measured while IR interventions were not in effect. Of these designs, 13 (93%) were scored as a 1 (consistent non-effects or non-therapeutic effects) while one was scored as a 2. No designs were identified with consistent positive generalized effects in the absence of intervention (i.e., scores of 3 or 4).

Only 13 (4%) designs included measurement of maintenance. Eight designs received a maintenance score of 1 (consistent non-effects or counter-therapeutic effects), two received scores of 3 (weak positive effects), and three received scores of 4 (strong positive effects). Five designs included maintenance measurement that occurred with ongoing treatment, six designs included within-session maintenance (e.g., behavior was measured immediately after IR interventions were withdrawn, in the same session), and only two included non-treatment sessions. Non-treatment sessions resulted in consistent non-effects (score: 1) for both designs, and within-session maintenance nearly always resulted in non-effects (score of 1 for five out of six designs, score of 3 for remaining design). See Table 7 and Figure 6.

Discussion

A relatively large body of published and unpublished research was identified, primarily evaluating vocal topographies of interruption and redirection (VIR, VIR+P), and primarily published in two journals, *JABA* or *Behavioral Interventions*. Evidence for publication bias was not found, which is inconsistent with one previous review (Dowdy et al., 2020). However, study quality was associated with outcomes, with fewer functional relations identified for less-rigorous studies. This indicates the potential value of including unpublished manuscripts in systematic reviews and assessing rigor of both published and unpublished work.

IR interventions were generally successful at reducing stereotypy but not for improving engagement or appropriate vocalizations. In addition, these interventions did not result in generalized or maintained behavior changes when the intervention was not in place. That is, all changes were *context-bound*, or occurring only when the intervention components were in place, often for only 5–10 min per session.

Contexts for intervention often included relatively sparse materials and interactions, with few reinforcement opportunities for engaging in appropriate behaviors. Although this may allow for increased experimental control, it also may serve as an establishing operation (Michael, 2000) for engaging in stereotypy and result in relatively more aversive interventions. Response allocation toward stereotypy may be high when other options are not available in baseline, essentially artificially inflating stereotypy when compared with typical environments (assuming that individuals have regular access to appropriate materials and interactions in their normal contexts). Relatively less is known about the effects of IR interventions when the individual has more response options that produce favorable outcomes across conditions, especially when those responses are incompatible with stereotypy (Dadds et al., 1988).

There were few studies conducted in the context of typical educational activities, although improved participation in these activities is often cited as a rationale for treating stereotypy. This is important because previous studies have found that effective teaching may result in reduction of stereotypy (Dib & Sturmey, 2013).

Signaled interventions resulted in more functional relations for motor stereotypy and appropriate vocalizations while unsignaled interventions resulted in more functional relations for vocal stereotypy. Perhaps participants were more likely to respond positively to these signals for motor stereotypy because signals may more effectively reduce unexpected physical contact,

since motor stereotypy more often resulted in physical interruptions, redirections, or error corrections. Additional work is needed to directly compare signaled and unsignaled intervention conditions.

Studies were often conducted in segregated settings, including separate school or residential settings for almost half of participants—this use of a punishment-based intervention may be concerning given the relative innocuousness of the target behaviors (i.e., few dangerous behaviors, mostly vocal stereotypy) and the potential vulnerability of people in segregated placements due to reduced observation of intervention by outsiders (cf. Morris & Hollins, 2021) and individual characteristics (e.g., inability to effectively self-advocate) that may be more prevalent in this population.

As has been pointed out in previous research (e.g., Wunderlich & Vollmer, 2015), VIR and VIR+P interventions often used interrupted measurement procedures, which could overestimate the effectiveness of these interventions, because they capture all stereotypy that occurs within baseline conditions, but do not capture stereotypy that occurs while IR is being actively implemented in intervention. We found evidence for this for vocal stereotypy, motor stereotypy, and appropriate vocalizations. The opposite pattern occurred for functional engagement, although there was less data available for this variable. It is possible that functional engagement measurement included prompted engagement, so including data during redirections (e.g., continuous measurement) may have resulted in more functional relations. More research is needed to determine the impact of measurement and including prompted versus independent engagement is needed. Overall, the impact of IR interventions on functional engagement may be somewhat difficult to interpret given both differences in continuous and interrupted measurement and the wide range of systems used (e.g., duration, momentary time sampling, partial interval

recording). This is a critical area of future research given increases in functional engagement may be necessary for long-term and generalized reductions in stereotypy,

IR Interventions as Evidence-Based?

Previous reviews have identified IR interventions as evidence-based, with at least one identifying numerous domains of improvement (Steinbrenner et al., 2020). Our review does not entirely agree with these past assessments. We did find the interventions to be useful for decreasing both vocal and motor stereotypy, but only when the intervention was actively being implemented. Importantly, we did not find evidence of generalized or maintained behavior change. Because the interventions are relatively intrusive, potentially aversive, and sometimes effortful, the lack of generalized behavior change is disappointing. Moreover, we found inconsistent outcomes for alternative behaviors like appropriate vocalizations and engagement. Thus, we identify IR interventions as evidence-based only for (a) context-bound behavior change (b) in the presence of continuous intervention implementation (c) in specialized settings.

Recommendations for Research

Results were generally mixed across studies; this may be partially explained by procedural differences across studies. When IR interventions are assessed via research, we offer the following four recommendations to improve validity and comparability of future research as well as social and ecological validity:

- 1) Avoid using baseline and intervention conditions that could inflate the likelihood of stereotypy (e.g., no materials and interactions; Dadds et al., 1988). Instead, use contexts that mirror those in which stereotypy is typically problematic. Consider that stereotypy in some contexts (e.g., free choice contexts like recess or leisure time when individuals can engage in any non-harmful behaviors they choose) should likely not be intervened on, or

be intervened on only with explicit assent of the participant, in the absence of overtly harmful effects (e.g., injury; Hagopian et al., 2011).

- 2) Because IR interventions can be high-effort for some participants (e.g., up to 94% of a session spent in treatment; Wunderlich et al., 2015) and potentially aversive, include information about participant assent and preference for procedures (Morris et al., 2021), and clearly articulate why the intervention is needed (Behavior Analyst Certification Board, 2020).
- 3) Use continuous measurement. That is, measure all occurrences of stereotypy in both baseline and intervention conditions. When measuring engagement and appropriate vocalizations, report separately independent responses and those prompted as part of redirections.
- 4) Measure engagement when evaluating interventions designed to decrease stereotypy. Use duration recording or momentary time sampling because other methods such as partial interval recording are poor estimates of duration (Ledford & Gast, 2018).
- 5) Conduct additional research is needed to determine when and if clinicians and educators should treat stereotypy and whether they can and will do so in typical contexts.
- 6) Conduct direct comparisons to evaluate whether certain components (e.g., signals, prompts for engagement) increase the likelihood of positive outcomes for IR interventions.

Recommendations for Practice

Based on the results of the review, it is clear that IR interventions are generally successful for reducing stereotypy while the intervention is in place, but that it does not result in generalized

changes in appropriate behaviors or when the intervention is withdrawn. Thus, we identify three recommendations for practitioners:

- 1) Because we have sufficient evidence that IR interventions do not result in generalized or maintained changes, only use IR interventions when context-bound behavior change is the goal. That is, do not use IR interventions when the treatment goal is reduction of stereotypy without continued use of IR interventions.
- 2) When IR interventions are used, explicitly inform endogenous caregivers and practitioners that the need for intervention procedures may be indefinite; when the intervention is withdrawn, the behaviors are likely to revert to baseline levels.
- 3) When IR interventions are required, do not use them in isolation. Instead, pair them with other interventions designed to support engagement and social interactions.

Limitations

The outcomes and recommendations derived from this review should be interpreted in light of several limitations. First, we did not calculate effect sizes or conduct a meta-analysis, which precluded the ability to determine whether *magnitude* of effect differed based on coded characteristics. We decided meta-analysis would be inappropriate given variability in measurement systems (e.g., duration, count, partial interval recording) that can cause differences in effect sizes (Pustejovsky, 2019) and potential differences caused by the use of continuous versus interrupted measurement systems. Second, we acknowledge that authors may not have anticipated improvements in secondary variables such as engagement and appropriate vocalizations, and that authors may have measured these variables to establish that no counter-therapeutic changes occurred. However, we consider it important to note that we found that IR interventions were not effective for improving these behaviors given this information does not

align with previous reviews (Steinbrenner et al., 2020). Finally, we did not analyze outcomes by intervention component (e.g., implementer responses to initiations or redirections). Given the wide variability in reported components (e.g., contextually appropriate responses to redirections only or to redirections and initiations, tangible reinforcers for engagement), each variation was present for only a few instances for each dependent variable. Thus, it was not reasonable to compare effectiveness across types.

Conclusions

Based on the results of this review, we identify IR interventions as effective for reducing stereotypy during intervention sessions but as ineffective for maintaining decreased stereotypy once intervention is removed, and also ineffective for increasing appropriate behaviors. Thus, we urge practitioners and researchers to use and evaluate these interventions only as part of a package that includes components designed to support active engagement and positive social interactions and to evaluate the maintenance and generality of these combined interventions in the absence of active treatment.

References

- Akers, J. S., Davis, T. N., Gerow, S., & Avery, S. (2020). Decreasing motor stereotypy in individuals with autism spectrum disorder: A systematic review. *Research in Autism Spectrum Disorders*, 77. <https://doi.org/10.1016/j.rasd.2020.101611>
- 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), 1-21. <https://doi.org/10.1002/jaba.30>
- Behavior Analyst Certification Board. (2020). *Ethics code for behavior analysts*. Littleton, CO: Author.
- Bury, S. M., Jellett, R., Spoor, J. R., & Hedley, D. (2020). “It defines who I am” or “It’s something I have”: What language do [autistic] Australian adults [on the autism spectrum] prefer?. *Journal of Autism and Developmental Disorders*, Advanced online publication. <https://doi.org/10.1007/s10803-020-04425-3>
- 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*, 30, 1210-1223. https://doi.org/10.1044/2020_AJSLP-20-00102
- Dadds, M., Schwartz, S., Adams, T., & Rose, S. (1988). The effects of social context and verbal skill on the stereotypic and task-involved behaviour of autistic children. *Journal of Child Psychology and Psychiatry*, 29, 669-676. <https://doi.org/10.1111/j.1469-7610.1988.tb01887.x>

- Dib, N., & Sturmey, P. (2007). Reducing student stereotypy by improving teachers' implementation of discrete-trial teaching. *Journal of Applied Behavior Analysis*, 40(2), 339-343. <https://doi.org/10.1901/jaba.2007.52-06>
- 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, 2151-2171. <https://doi.org/10.1002/jaba.724>
- Gibbs, A. R., & Tullis, C. A. (2020). The Emergence of Untrained Relations in Individuals with Autism and Other Intellectual and Developmental Disabilities: a Systematic Review of the Recent Literature. *Review Journal of Autism and Developmental Disorders*, 1-26. <https://doi.org/10.1007/s40489-020-00211-0>
- Goldman, S., Wang, C., Salgado, M. W., Greene, P. E., Kim, M., & Rapin, I. (2009). Motor stereotypies in children with autism and other developmental disorders. *Developmental Medicine & Child Neurology*, 51, 30-38. <https://doi.org/10.1111/j.1469-8749.2008.03178.x>
- Haddaway, N. R., McGuinness, L. A., & Pritchard, C. C. (2021). PRISMA2020: R package and ShinyApp for producing PRISMA 2020 compliant flow diagrams (Version 0.0.2).
- Hagopian, L. P., González, M. L., Rivet, T. T., Triggs, M., & Clark, S. B. (2011). Response interruption and differential reinforcement of alternative behavior for the treatment of pica. *Behavioral Interventions*, 26(4), 309-325. <https://doi.org/10.1002/bin.339>
- Heathers, J. A., Gilchrist, K. H., Hegarty-Craver, M., Grego, S., & Goodwin, M. S. (2019). An analysis of stereotypical motor movements and cardiovascular coupling in individuals on the autism spectrum. *Biological Psychology*, 142, 90-99. <https://doi.org/10.1016/j.biopsycho.2019.01.004>

- Joyce, C., Honey, E., Leekam, S. R., Barrett, S. L., & Rodgers, J. (2017). Anxiety, intolerance of uncertainty and restricted and repetitive behaviour: Insights directly from young people with ASD. *Journal of Autism and Developmental Disorders*, 47, 3789-3802.
<https://doi.org/10.1007/s10803-017-3027-2>
- Kenny, L., Hattersley, C., Molins, B., Buckley, C., Povey, C., & Pellicano, E. (2016). Which terms should be used to describe autism? Perspectives from the UK autism community. *Autism*, 20, 442-462. <https://doi.org/10.1177/1362361315588200>
- 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, 1234-1243. <https://doi.org/10.1016/j.rasd.2013.07.009>
- Leaf, J. B., Sato, S. K., Javed, A., Arthur, S. M., Creem, A. N., Cihon, J. H., ... & Oppenheim-Leaf, M. L. (2021). The evidence-based practices for children, youth, and young adults with autism report: Concerns and critiques. *Behavioral Interventions*, 36, 457-472.
<https://doi.org/10.1002/bin.1771>
- Ledford, J. R., Chazin, K. T., Lane, J. D., Zimmerman, K. N., & Ayres, K. A. (2020, September). Single case analysis and review framework (SCARF). Retrieved from: <http://ebip.vkcsites.org/scarfv2>
- Ledford, J. R., & Gast, D. L. (Eds.). (2014). *Single case research methodology: Applications in special education and behavioral sciences*. Routledge.
- Ledford, J. R., Lambert, J. M., Barton, E. E., & Ayres, K. M. (2021). The Evidence Base for Interventions for Individuals With ASD: A Call to Improve Practice Conceptualization and Synthesis. *Focus on Autism and Other Developmental Disabilities*, 10883576211023349.

- Ledford, J. R., & Windsor, S. A. (2020). 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>
- Lydon, S., Healy, O., O'Reilly, M., & McCoy, A. (2013). A systematic review and evaluation of response redirection as a treatment for challenging behavior in individuals with developmental disabilities. *Research in Developmental Disabilities*, 34, 3148-3158. <https://doi.org/10.1016/j.ridd.2013.06.010>
- Matson, J. L., Kiely, S. L., & Bamburg, J. W. (1997). The effect of stereotypes on adaptive skills as assessed with the DASH-II and Vineland Adaptive Behavior Scales. *Research in Developmental Disabilities*, 18, 471-476. [https://doi.org/10.1016/S0891-4222\(97\)00023-1](https://doi.org/10.1016/S0891-4222(97)00023-1)
- Manor-Binyamini, I., & Schreiber-Divon, M. (2019). Repetitive behaviors: Listening to the voice of people with high-functioning autism spectrum disorder. *Research in Autism Spectrum Disorders*, 64, 23-30. <https://doi.org/10.1016/j.rasd.2019.04.001>
- Michael, J. (2000). Implications and refinements of the establishing operation concept. *Journal of Applied Behavior Analysis*, 33(4), 401-410. <https://doi.org/10.1901/jaba.2000.33-401>
- Morris, C., & Hollins, N. A. (2021). On the Uncanny Similarities Between Police Brutality and Client Mistreatment. *Behavior Analysis in Practice*, 1-6. <https://doi.org/10.1007/s40617-021-00576-6>
- Morrison, K., & Rosales-Ruiz, J. (1997). The effect of object preferences on task performance and stereotypy in a child with autism. *Research in Developmental Disabilities*, 18, 127-137. [https://doi.org/10.1016/S0891-4222\(96\)00046-7](https://doi.org/10.1016/S0891-4222(96)00046-7)

- Ouzzani, M., Hammady, H. Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan — a web and mobile app for systematic reviews. *Systematic Reviews*, 5, 210-220.
<https://doi.org/10.1186/s13643-016-0384-4>
- Powell, J. L., Pringle, L., & Greig, M. (2017). Investigation of the association between motor stereotypy behavior with fundamental movement skills, adaptive functioning, and autistic spectrum disorder symptomology in children with intellectual disabilities. *Journal of Child Neurology*, 32, 222-229. <https://doi.org/10.1177/0883073816678551>
- Pustejovsky, J. E. (2019). Procedural sensitivities of effect sizes for single-case designs with directly observed behavioral outcome measures. *Psychological Methods*, 24, 217. <https://doi.org/10.1037/met0000179>
- Sherer, M. R., & Schreibman, L. (2005). Individual behavioral profiles and predictors of treatment effectiveness for children with autism. *Journal of Consulting and Clinical Psychology*, 73, 525–538.
- Spencer, V. G., & Alkhanji, R. (2018). Response interruption and redirection (RIRD) as a behavioral intervention for vocal stereotypy: A systematic review. *Education and Training in Autism and Developmental Disabilities*, 53, 33-43.
<https://www.jstor.org/stable/26420425>
- Steinbrenner, J. R., Hume, K., Odom, S. L., Morin, K. L., Nowell, S. W., Tomaszewski, B., ... & Savage, M. N. (2020). Evidence-Based Practices for Children, Youth, and Young Adults with Autism. *FPG Child Development Institute*.
- What Works Clearinghouse (2020). *Procedures and standards handbook*. Version 4.1. Author. Retrieved from: <https://ies.ed.gov/ncee/wwc/Docs/referenceresources/WWC-Standards-Handbook-v4-1-508.pdf>

- Wunderlich, K. L., & Vollmer, T. R. (2015). Data analysis of response interruption and redirection as a treatment for vocal stereotypy. *Journal of Applied Behavior Analysis*, 48, 749-764. <https://doi.org/10.1002/jaba.227>
- Zimmerman, K. N., Ledford, J. R., Severini, K. E., Pustejovsky, J. E., Barton, E. E., & Lloyd, B. P. (2018). Single-case synthesis tools I: Comparing tools to evaluate SCD quality and rigor. *Research in Developmental Disabilities*, 79, 19-32. <https://doi.org/10.1016/j.ridd.2018.02.003>

Table 1. Dependent variables measured by intervention type.

	PIR	VIR	VIR + P	Mixed	<i>Total</i>
Vocal Stereotypy	3	66	65	0	<i>134</i>
Motor Stereotypy	3	24	14	11 ^a	<i>52</i>
Appropriate Vocalizations	0	38	5	1	<i>44</i>
Functional Engagement	3	11	14	5	<i>33</i>
Combined Vocal/Motor Stereotypy	0	7	3	3	<i>13</i>
Dangerous Behaviors	3	1	3	0	<i>7</i>
Task Completion	0	2	0	2	<i>4</i>
Other	1	2	1	0	<i>4</i>
Total	13	151	105	22	<i>291</i>

Note: PIR = physical interruption and redirection. VIR = vocal interruption and redirection.

VIR+P = vocal interruption and redirection with physical prompting for error correction. Mixed = interruption and redirection intervention that does not fit any other category; a = includes two designs assessing hoarding, ordering, and arranging.

Table 2. Contextual relevance of redirection.

	PIR	VIR	VIR + P	Mixed	<i>Total</i>
Contextually relevant	12	15	13	15	<i>55</i>
Contextually irrelevant	0	101	53	4	<i>158</i>
No appropriate context	1	33	20	3	<i>57</i>
Unclear/not reported	0	2	19	0	<i>21</i>

Note: PIR = physical interruption and redirection. VIR = vocal interruption and redirection.

VIR+P = vocal interruption and redirection with physical prompting for error correction. Mixed = interruption and redirection intervention that does not fit any other category.

Table 3. Intervention signals and stop criteria.

	PIR	VIR	VIR + P	Mixed	<i>Total</i>
Signal used					
Signaled intervention	5	58	50	7	<i>120</i>
Un-signaled interventions	8	93	55	15	<i>171</i>
Stop criterion used					
Presentation of Demands	0	112	0	2	<i>114</i>
Compliance with Demands	0	20	99	5	<i>124</i>
No criterion reported	13	19	6	15	<i>53</i>

Note: PIR = physical interruption and redirection. VIR = vocal interruption and redirection.

VIR+P = vocal interruption and redirection with physical prompting for error correction. Mixed = interruption and redirection intervention that does not fit any other category.

Table 4. Visual analysis outcomes by study quality.

	Lowest			Highest
	quartile	2 nd quartile	3 rd quartile	quartile
	(0.97-1.86)	(1.86-2.22)	(2.22-2.47)	(2.53-3.36)
Functional relation, no weak effects (4)	25 (35%)	34 (47%)	36 (50%)	40 (56%)
Functional relation, with weak effect (3)	18 (25%)	11 (15%)	17 (24%)	14 (19%)
Inconsistent effects (2)	1 (1%)	1 (1%)	0 (0%)	2 (3%)
Inconsistent effects (1)	8 (11%)	3 (4%)	9 (13%)	2 (3%)
Non-effects of negative effects (0)	20 (28%)	23 (32%)	10 (14%)	14 (19%)
Total Functional Relations	43 (60%)	45 (63%)	53 (74%)	54 (75%)

Note: Numbers 0–4 following description of outcomes indicate scores on Single Case Analysis and Review Framework (SCARF; Ledford et al., 2020). Functional relations were identified for scores of 3 and 4 on the SCARF. Each quartile represents 72 scores, with scores from one study between each quartile (total of 3) discarded for this analysis.

Table 5. Visual analysis outcomes by type.

	PIR	VIR	VIR + P	Mixed	<i>Total</i>
Functional relation, no weak effects (4)	7	62	49	17	135 (46%)
Functional relation, with weak effect (3)	2	41	18	1	62 (21%)
Inconsistent effects (2)	0	2	2	0	4 (1%)
Inconsistent effects (1)	0	14	9	0	23 (8%)
Non-effects of negative effects (0)	4	32	27	4	67 (23%)
Total Functional Relations	9	103	68	18	197
	(69%)	(68%)	(65%)	(82%)	(68%)

Note: Numbers 0–4 following description of outcomes indicate scores on Single Case Analysis and Review Framework (SCARF; Ledford et al., 2020). PIR = physical interruption and redirection. VIR = vocal interruption and redirection. VIR+P = vocal interruption and redirection with physical prompting for error correction. Mixed = interruption and redirection intervention that does not fit any other category.

Table 6. Functional Relations by dependent variable and intervention type

	PIR	VIR	VIR + P	Mixed	Total
Vocal Stereotypy	1/3 (33%)	56/66 (85%)	43/65 (66%)	--	100/131 (75%)
Motor Stereotypy	2/3 (67%)	22/24 (92%)	13/14 (93%)	10/11 (91%)	47/52 (90%)
Appropriate Vocalizations	--	12/38 (32%)	0/5 (0%)	--	12/43 (28%)
Functional Engagement	2/3 (67%)	5/11 (45%)	6/14 (43%)	5/5 (100%)	18/33 (55%)

Note: These data represent the number of designs with functional relations (visual analysis scores of 3 or 4 on the SCARF). Data only included if number of designs included in the cell is 3 or more. Note: PIR = physical interruption and redirection. VIR = vocal interruption and redirection. VIR+P = vocal interruption and redirection with physical prompting for error correction. Mixed = interruption and redirection intervention that does not fit any other category.

Table 7. Positive maintenance and generalization outcomes by intervention type.

	PIR	VIR	VIR + P	Mixed	<i>Total</i>
<i>Treatment withdrawn</i>					
Generalization	--	0/6	0/5	0/3	0/14
		0%	0%	0%	0%
Maintenance	--	0/1	0/1	--	0/2
Across sessions		0%	0%		0%
Maintenance	--	0/2	1/4	--	1/6
Within session		0%	25%		17%
<i>Treatment in place</i>					
Generalization	2/2	6/9	5/9	2/2	15/22
	100%	67%	56%	100%	68%
Maintenance	1/1	2/2	1/2	--	4/5
	100%	100%	50%		80%

Note: Positive outcomes refer to consistent positive effects (scores of 3 or 4 on SCARF). VIR = vocal interruption and redirection. VIR+P = vocal interruption and redirection with physical prompting for error correction. Mixed = interruption and redirection intervention that does not fit any other category.

Figure 1

Visual analysis outcomes for each intervention type.

Note: PIR = physical interruption and redirection. VIR = vocal interruption and redirection. VIR+P = vocal interruption and redirection with physical error correction, when needed. Mixed = interruption and redirection interventions that do not qualify for any of the other categories.

Figure 2

Visual analysis outcomes for physical interruption and redirection (PIR) interventions, across dependent variables.

Note: No appropriate vocalization outcomes were measured for PIR interventions.

Figure 3

Visual analysis outcomes for vocal interruption and redirection (VIR) interventions, across dependent variables.

Figure 4
Visual analysis outcomes for vocal interruption and redirection with physical error correction (VIR+P) interventions, across dependent variables.

Figure 5

Visual analysis outcomes for mixed interruption and redirection interventions, across dependent variables.

Supplemental Materials Tables and Figures

Agreement information.

Category	Agreement Percentage	Range across Codes	Percentage of Sample Coded
Screening			
Database search (n = 399)	94%	n/a	100%
Other methods (n = 95)	94%	n/a	100%
Full text review (n = 126)	89%	n/a	100%
Descriptive Coding			
Source-Level Codes (n = 53)	100%	n/a	100%
Participant-Level Codes (n = 133)			
Dependent variables	94%	83-100%	100%
Participants and contexts	94%	75-100%	100%
Design-Level Codes (n = 291)			
Designs included	98%	92-100%	100%
Dependent variables	95%	83-100%	100%
Baseline	94%	91-98%	100%
Intervention	95%	82-100%	100%
SCARF Coding (n = 291)			
Rigor (Internal validity)	95%	90-98%	55%
Quality (external validity)	95%	87-97%	55%
Outcomes	92%	83-100%	55%

Note: Agreement calculated as agreements divided by agreements plus disagreements. See supplemental materials for individual codes included in each category. All discrepancies were reviewed, and the final code determined via consensus; in case of disagreement, the first author was consulted/determined the final code. SCARF = Single Case Analysis and Review Framework (Ledford et al., 2020). The category with lower than 80% agreement was occurred when coders did not agree on “other” reported setting type.

Search terms according to PICOS structure.

Domain	Terms
Participants	(asd or autism or pdd or Asperger) AND
Intervention	("response NEAR/2 interruption" OR "RIRD" OR "response blocking" or "abolishing operation") AND
Comparisons	N/A
Outcomes	("challenging behavior" or stereotyp* or self-injury or pica or rumination or bruxism or "automatically maintained")
Study Type	N/A

Note: N/A indicates no search terms were entered for comparisons or study type. Each of three search strings were combined with the AND Boolean operator.

Intervention definitions

Type	Definition	Examples
Vocal interruption and redirection (VIR)	Vocal interruption is followed by vocal redirection.	<i>Compliance with redirection</i> Participant: <engages in stereotypy> Implementer interruption: “Georgia” Implementer redirection: “Touch your nose like this” <demonstrates> Participant: <complies> END OF TRIAL <i>Noncompliance with redirection</i> Participant: <engages in stereotypy> Implementer interruption: “Georgia” Implementer redirection: “Touch your nose like this” <demonstrates> Participant: <does not touch nose> Implementer response: “Touch your nose like this” <demonstrates> Participant: <complies> END OF TRIAL
Vocal interruption and redirection with physical error correction (VIR+P)	VIR sometimes includes subsequent physical guidance (i.e., vocal interruption, followed by least-to-most prompting of the redirection)	<i>Compliance with redirection</i> Same as VIR <i>Noncompliance with redirection</i> Participant: <engages in stereotypy> Implementer interruption: “Georgia” Implementer redirection: “Touch your nose like this” <demonstrates> Participant: <does not touch nose> Implementer response: “Touch your nose like this” <hand over hand guidance> Participant: <complies> END OF TRIAL
Physical interruption and redirection (PIR)	Physical interruption is followed by physical redirection	Participant: <engages in stereotypy> Implementer interruption: <places hands on table to block movement> Implementer redirection: <hand over hand guidance to play with toys> Participant: <complies> END OF TRIAL
Mixed interruption and redirection intervention (Mixed)	Interventions including both interruption and redirection which do not meet criteria for the three categories above.	<i>Vocal interruption, physical redirection</i> Participant: <engages in stereotypy> Implementer interruption: “Georgia” Implementer redirection: <hand over hand guidance to play with toys> Participant: <complies> END OF TRIAL <i>Physical interruption, vocal redirection</i> Participant: <engages in stereotypy> Implementer interruption: <places hands on table to block movement> Implementer redirection: “play with the toy like this” <demonstrates> Participant: <complies> END OF TRIAL

Dependent variables measured by recording system.

	Duration	Partial	Count	Whole	Other
Vocal Stereotypy	99	24	6	0	5
Motor Stereotypy	28	18	0	1	3
Appropriate Vocalizations	2	11	31	0	0
Functional Engagement	15	8	6	4	0
Combined Vocal/Motor Stereotypy	8	4	0	1	0
Dangerous Behaviors	1	1	5	0	0
Task Completion	4	0	0	0	0
Other	1	1	2	2	0
Total	158	67	50	8	8

Note: Partial = partial interval recording. Whole = whole interval recording.

Topography of target behaviors and redirected behaviors.

	PIR	VIR	VIR + P	Mixed	Total
Target behaviors					
Self-injurious behavior	1	0	6	0	7
Vocal behavior	2	111	70	15	198
Motor behavior	10	34	24	0	68
Both vocal and motor behavior	0	6	5	7	18
Redirected behaviors					
Vocal response	0	84	28	0	112
Motor response	11	26	70	17	124
Both motor and vocal response	0	16	0	0	16
Not reported/unclear	2	25	7	5	39
Topographical match between target and redirected behaviors					
Matched	11	101	55	15	182
Unmatched	0	11	42	0	53
Partial match	0	14	1	2	17
Unclear	2	25	7	5	39

Note: IR procedures were conducted contingent on target behaviors. Redirected behaviors were those expected to replace the target behaviors. Partially matched behaviors were those in which both motor and vocal were targeted but the redirected behavior was motor or vocal, or in which either motor or vocal were targeted but the redirected behavior was both. PIR = physical interruption and redirection. VIR = vocal interruption and redirection. VIR+P = vocal interruption and redirection with physical prompting for error correction. Mixed = interruption and redirection intervention that does not fit any other category.

Visual analysis outcomes by publication status.

	Published (n = 200, 69%)	Unpublished (n = 91, 31%)	Total
Functional relation, no weak effects (4)	92 (46%)	43 (47%)	135 (46%)
Functional relation, with weak effect (3)	44 (22%)	18 (20%)	62 (21%)
Inconsistent effects (2)	3 (2%)	1 (1%)	4 (1%)
Inconsistent effects (1)	18 (9%)	5 (5%)	23 (8%)
Non-effects of negative effects (0)	43 (22%)	24 (26%)	67 (23%)
Total Functional Relations	136 (68%)	61 (67%)	197 (68%)

Note: Numbers 0–4 following description of outcomes indicate scores on Single Case Analysis and Review Framework (SCARF; Ledford et al., 2020). Functional relations were identified for scores of 3 and 4 on the SCARF.

Functional relations by dependent variable and measurement type for VIR and VIR+P interventions

	I	C
Vocal Stereotypy	55/66 (83%)	34/48 (71%)
Motor Stereotypy	20/20 (100%)	15/18 (83%)
Appropriate Vocalizations	11/19 (58%)	0/19 (0%)
Functional Engagement	6/14 (43%)	5/10 (50%)

Note: These data represent the number of designs with functional relations (visual analysis scores of 3 or 4 on the SCARF). I = interrupted measurement. C = continuous measurement.

Functional relations by dependent variable signal status.

	Signal	No Signal
Vocal Stereotypy	42/67 (63%)	58/64 (91%)
Motor Stereotypy	25/25 (100%)	24/29 (83%)
Appropriate Vocalizations	4/8 (50%)	6/34 (18%)

PRISMA diagram for search and screening procedures.

Note: Diagram generated via PRISMA2020 (Haddaway et al., 2021).

Visual analysis outcomes for all designs, designs in published sources, and designs in unpublished sources.

Note: Outcomes of 3 and 4 are consistent with functional relations. Designs with higher quality and rigor are depicted further right on the graph.

Supplemental References

Sources Included in the Systematic Review

- Abrena, I. L. (2016). Further evaluation of the effects of response interruption and redirection on vocal stereotypy in children with autism. Unpublished Doctoral Dissertation. The Chicago School of Professional Psychology.
- Ahearn, W. H., Clark, K. M., MacDonald, R. P., & Chung, B. I. (2007). Assessing and treating vocal stereotypy in children with autism. *Journal of Applied Behavior Analysis, 40*(2), 263-275.
- Ahrens, E. N., Lerman, D. C., Kodak, T., Worsdell, A. S., & Keegan, C. (2011). Further evaluation of response interruption and redirection as treatment for stereotypy. *Journal of Applied Behavior Analysis, 44*(1), 95-108.
- Andrejczyk, L. (2010). Evaluating the use of a multiple schedule as a method for identifying treatment effects. Unpublished Master's Thesis. Northeastern University.
- Arabian, A. (2013). Effects of response interruption and redirection in naturalistic settings. Unpublished Master's Thesis. California State University, Los Angeles.
- Bharucha, D. (2011). Evaluating the use of a multiple schedule for identifying treatment and motivational effects. Unpublished Doctoral Dissertation. Northeastern University.
- Brusa, E., & Richman, D. (2008). Developing stimulus control for occurrences of stereotypy exhibited by a child with autism. *International Journal of Behavioral Consultation and Therapy, 4*, 264-269.
- Cannella, H. I. (2005). Assessment and treatment of automatically maintained hand mouthing in individuals with developmental disabilities. Unpublished Doctoral Dissertation. University of Texas at Austin.
- Carroll, R. A., & Kodak, T. (2014). An evaluation of interrupted and uninterrupted measurement of vocal stereotypy on perceived treatment outcomes. *Journal of Applied Behavior Analysis, 47*(2), 264-276.
- Cassella, M. D., Sidener, T. M., Sidener, D. W., & Progar, P. R. (2011). Response interruption and redirection for vocal stereotypy in children with autism: A systematic replication. *Journal of Applied Behavior Analysis, 44*(1), 169-173.
- Cividini-Motta, I., Garcia, A. R., Livingston, I., & MacNaul, H. L. (2019). The effect of response interruption and redirection with and without a differential reinforcement of alternative behavior component on stereotypy and appropriate responses. *Behavioral Interventions, 34*(1), 3-18.
- Colón, I. L., & Ahearn, W. H. (2019). An analysis of treatment integrity of response interruption and redirection. *Journal of Applied Behavior Analysis, 52*(2), 337-354.
- DeRosa, N. M., Novak, M. D., Morley, A. J., & Roane, H. S. (2019). Comparing response blocking and response interruption/redirection on levels of motor stereotypy: Effects of data analysis procedures. *Journal of Applied Behavior Analysis, 52*(4), 1021-1033.
- Duker, P. I., & Schappveld, M. (1996). Increasing on-task behaviour through interruption-prompting. *Journal of Intellectual Disability Research, 40*(4), 291-297.
<https://doi.org/10.1111/j.1365-2788.1996.tb00633.x>
- Falligant, J. M., & Dommetrup, A. K. (2020). Decreasing motor stereotypy in leisure contexts: Analysis of stimulus control with response interruption and redirection. *Behavioral Interventions, 35*(2), 306-314.

- Farber, R. (2010). The effects of response interruption redirection and differential reinforcement of other behavior on motor stereotypy. Unpublished Master's Thesis. Northeastern University.
- Gartland, M. E. (2009). The effects of response interruption/redirection and differential reinforcement of other behaviors on vocal stereotypy in children with autism. Unpublished master's thesis, The Ohio State University
- Gauthier, K. A., Ahearn, W. H., & Colón, I. L. (2020). Further evaluation of treatment integrity for response interruption and redirection. *Behavioral Interventions*, 35(4), 571-580.
- Gavidia, G. (2010). Comparison of reinforcement alone and in comparison with redirection for treating automatically-reinforced motor stereotypy. Unpublished Master's Thesis. Northeastern University..
- Gerow, S., Rivera, G., Akers, J. S., Kirkpatrick, M., & Radhakrishnan, S. (2019). Parent-implemented treatment for automatically maintained stereotypy. *Behavioral Interventions*, 34(4), 466-474.
- Gilbride, M. M. (2011). A comparison of differential reinforcement of other behaviors and response interruption and redirection on vocal stereotypy. Unpublished Doctoral Dissertation. Northeastern University.
- Giles, A. F., Peter, I. I. S., Penle, S. T., & Gibson, A. B. (2012). Preference for blocking or response redirection during stereotypy treatment. *Research in Developmental Disabilities*, 33(6), 1691-1700.
- Giles, A., Swain, S., Quinn, L., & Weifenbach, B. (2018). Teacher-implemented response interruption and redirection: Training, evaluation, and descriptive analysis of treatment integrity. *Behavior Modification*, 42(1), 148-169.
- Gould, K. M., Harper, J. M., Gillich, E., & Luiselli, J. K. (2019). Intervention, stimulus control, and generalization effects of response interruption and redirection on motor stereotypy. *Behavioral Interventions*, 34(1), 118-126.
- Hagopian, L. P., & Toole, L. M. (2009). Effects of response blocking and competing stimuli on stereotypic behavior. *Behavioral Interventions*, 24(2), 117-125.
- Hagopian, L. P., González, M. L., Rivet, T. T., Triggs, M., & Clark, S. B. (2011). Response interruption and differential reinforcement of alternative behavior for the treatment of pica. *Behavioral Interventions*, 26(4), 309-325. <https://doi.org/10.1002/bin>
- Jared, C. E. (2011). The effects of response interruption redirection and differential reinforcement of other behaviors on rates of vocal stereotypy. Unpublished master's thesis. The Ohio State University.
- Joung, E. G. (2011). The effects of response interruption and redirection. Unpublished Doctoral Dissertation. University of Kansas.
- King, A. D. (2011). A Model for Determining an Effective and Least Restrictive Treatment to Reduce Vocal Stereotypy. Unpublished Master's Thesis. University of Wisconsin.
- Korneder, J. (2014). The use of response interruption redirection, time-out, and differential reinforcement to decrease stereotypy. Unpublished Doctoral Dissertation. Western Michigan University.
- Liu-Gitz, L., & Banda, D. R. (2010). A replication of the RIRD strategy to decrease vocal stereotypy in a student with autism. *Behavioral Interventions*, 25(1), 77-87.
- Long, R. M. (2015). Response interruption and redirection applied to life skills tasks. Unpublished Doctoral Dissertation. University of Maine.

- Love, J. J., Miguel, C. F., Fernand, J. K., & LaBrie, J. K. (2012). The effects of matched stimulation and response interruption and redirection on vocal stereotypy. *Journal of Applied Behavior Analysis*, 45(3), 549-564.
- Martinez, C. K. (2014). Evaluating the efficacy of response interruption and redirection (RIRD) Using empirically-derived consequences. Unpublished Doctoral Dissertation. Florida Institute of Technology.
- Martinez, C. K., Betz, A. M., Liddon, I. J., & Werle, R. L. (2016). A progression to transfer RIRD to the natural environment. *Behavioral Interventions*, 31(2), 144-162.
- McNamara, K., & Cividini-Motta, I. (2019). Further evaluation of treatments for vocal stereotypy: Response interruption and redirection and response cost. *Behavioral Interventions*, 34(2), 181-197.
- Miguel, I. F., Clark, K., Tereshko, L., Ahearn, W. H., & Zarcone, J. (2009). The effects of response interruption and redirection and sertraline on vocal stereotypy. *Journal of Applied Behavior Analysis*, 42(4), 883-888.
- Muething, C., Call, N. A., & Clark, S. (2020). An evaluation of differential reinforcement in the treatment of pica. *Developmental Neuropsychology*, 23(5), 321-327.
- Pastrana, S. J., Rapp, J. T., & Frewing, T. M. (2013). Immediate and subsequent effects of response interruption and redirection on targeted and untargeted forms of stereotypy. *Behavior Modification*, 37(4), 591-610.
- Peña, K. (2017). A component analysis of response interruption and redirection for vocal stereotypy in children with autism spectrum disorder. Unpublished Master's Thesis. University of South Florida.
- Perry, T. L. (2012). The effects of response interruption and redirection on language skills in children with vocal stereotypy. Unpublished Doctoral Dissertation. Eastern Michigan University.
- Plavnick, J. B., Mariage, T., Englert, I. S., Constantine, K., Morin, L., & Skibbe, L. (2014). Promoting independence during computer assisted reading instruction for children with autism spectrum disorders. *Revista Mexicana de Analisis de la Conducta*, 40(2), 85-105. <https://doi.org/10.5514/rma1.v40.i2.63667>
- Saini, V., Gregory, M. K., Uran, K. J., & Fantetti, M. A. (2015). Parametric analysis of response interruption and redirection as treatment for stereotypy. *Journal of Applied Behavior Analysis*, 48(1), 96-106.
- Schlichenmeyer, K. (2011). Evaluating the generalization and maintenance of treatment effects through stimulus control based interventions. Unpublished Master's Thesis. Northeastern University.
- Schmidt, J. D., Long, A., Goetzel, A. L., Tung, C., Pizarro, E., Phillips, C., & Hausman, N. (2017). Decreasing pica attempts by manipulating the environment to support prosocial behavior. *Journal of Developmental and Physical Disability*, 29(5), 683-697. <https://doi.org/10.1007/s10882-017-9548-y>
- Shawler, L. A., & Miguel, I. F. (2015). The effects of motor and vocal response interruption and redirection on vocal stereotypy and appropriate vocalizations. *Behavioral Interventions*, 30(2), 112-134.
- Shawler, L. A., Dianda, M., & Miguel, C. F. (2020). A comparison of response interruption and redirection and competing items on vocal stereotypy and appropriate vocalizations. *Journal of Applied Behavior Analysis*, 53(1), 355-365.

- Spector, V. (2018). Behavioral treatment of stereotypy in children with autism spectrum disorder: a comparative analysis. Unpublished Doctoral Dissertation. The Claremont Graduate University.
- Steinhauser, H. M. (2020). Examining stereotypy when promoting appropriate behavior: Reinforcement and response interruption and redirection. Unpublished Doctoral Dissertation. Western New England University.
- Toper-Korkmaz, O., Lerman, D. C., & Tsami, L. (2018). Effects of toy removal and number of demands on vocal stereotypy during response interruption and redirection. *Journal of Applied Behavior Analysis*, 51(4), 757-768.
- Verriden, A. L., & Roscoe, E. M. (2019). An evaluation of a punisher assessment for decreasing automatically reinforced problem behavior. *Journal of Applied Behavior Analysis*, 52(1), 205-226.
- Wells, J., Collier, L., & Sheehey, P. H. (2016). Decreasing vocal stereotypy of a youth with autism in a classroom setting. *Child & Family Behavior Therapy*, 38(2), 164-174.
- Wunderlich, K. L., & Vollmer, T. R. (2015). Data analysis of response interruption and redirection as a treatment for vocal stereotypy. *Journal of Applied Behavior Analysis*, 48(4), 749-764.