A Meta-Analysis of Single-Case Research on Applied Behavior Analytic Interventions for Individuals with Down Syndrome

Nicole Neil, Ashley Amicarelli, Brianna Anderson, Kailee Liesemer

Corresponding Author:
Nicole Neil, Ph.D., BCBA-D, Assistant Professor
nneil@uwo.ca
Western University
Faculty of Education
1137 Western Road
London, On, N6G 1G7, Canada

Ashley Rose Amicarelli, M.A.
aamicare@uwo.ca
Western University
Faculty of Education
1137 Western Road
London, On, N6G 1G7, Canada

Brianna Michelle Anderson, M.A., BCBA
bsteeper@uwo.ca
Western University
Faculty of Education
1137 Western Road
London, On, N6G 1G7, Canada

Kailee Liesemer, B.A.
klieseme@uwo.ca
Western University
Faculty of Education
1137 Western Road
London, On, N6G 1G7, Canada

Abstract

This systematic review (Registration number: CRD42018089688) evaluates single-case research design studies investigating applied behavior analytic (ABA) interventions for individuals with Down syndrome (DS). One-hundred-twenty-five studies examining the efficacy of ABA interventions on increasing skills and/or decreasing challenging behaviors met inclusion criteria. The What Works Clearinghouse standards and Risk of Bias in N-of-1 Trials scale were used to analyze methodological characteristics, and Tau-U effect sizes were calculated. Results suggest the use of ABA-based interventions are promising for behavior change in individuals with DS. Thirty-six high quality studies were identified and demonstrated a medium overall effect. A range of outcomes was targeted, primarily communication and challenging behavior. These outcomes will guide future research on ABA interventions and DS.

Keywords: Down syndrome, single-case research, Tau-U, applied behavior analysis

A Meta-Analysis of Single-Case Research on Applied Behavior Analytic Interventions for Individuals with Down Syndrome

Down syndrome (DS), the most common genetic cause of intellectual disability, is a result of the presence of an extra chromosome 21 at birth and affects 1 in 1000 births worldwide (World Health Organization [WHO], 2017). Individuals with DS show an early developmental pattern of strengths and weaknesses, termed the DS behavioral phenotype, with developmental delays in physical, behavioral, communicative, cognitive and/or social domains (Fidler, 2005). Many individuals with DS require lifelong interventions to address challenging behaviors and facilitate engagement in everyday activities in ways that enhance their overall quality of life.

The beneficial effects of various interventions for individuals with DS and their families have been documented in the literature. Interventions have been shown to improve challenging behaviors (Cole & Levinson, 2002; Feeley & Jones, 2006), motor development (Ulrich, Ulrich, Angulu-Kinzler, &Yun, 2001), phonological awareness (Kennedy & Flynn, 2003), school inclusion (Buckley, Bird, Sacks, & Archer 2006), adaptive functioning (Jacola, Hickey, Howe, Esbensen, & Shear, 2014) and quality of life (Crook, Adams, Shorten, & Langdon, 2016). Despite the variety of interventions targeting weaknesses identified in the DS behavioral phenotype, it is notable that several reviews identify applied behavior analytic (ABA) approaches such as prompts and reinforcement as effective for in addressing multiple areas of weakness identified in the DS behavioral phenotype (Feeley & Jones, 2006; Neil & Jones, 2018; New York State Department of Health [NYSDOH], 2006). ABA is the application of the principles of learning (e.g., reinforcement, extinction) to areas of social significance (Baer, Wolf, & Risley, 1968). ABA interventions are characterized by data-based decision making and

empirical demonstrations of a functional relationship between changes in behavior and the intervention being implemented (Baer et al., 1968).

Prior Reviews of Applied Behavior Analysis Applications in Down Syndrome

In 2006, NYSDOH reviewed single-case and group studies to identify effective interventions targeting multiple areas for young children with DS. Interventions were categorized by strength of evidence; an intervention was considered to have strong evidence when they had two or more high-quality studies supporting their use. The NYSDOH guidelines recommended intervention strategies use the principles of ABA such as prompting and positive reinforcement. Although this study reviewed the level of evidence for various interventions for DS, it only reported summary ratings for the intervention strategies (strong to limited evidence) and it did not quantify the effects associated with their use.

In the only meta-analysis to date, Neil and Jones (2018) systematically reviewed 37 studies of intervention targeting various communication skills in individuals with DS. They found the majority of studies (78.3%) used behavioral approaches including prompting and reinforcement and were associated with positive outcomes. Neil and Jones (2018) also recommended behavioral approaches to intervention. This study reported on methodological components of the studies (e.g., interobserver agreement and intervention integrity) but did not include a rating or the quality of the evidence. O'Toole et al. evaluated randomized controlled trials of parent-mediated interventions for communication in young children with DS including three studies, but no quantitative synthesis due to insufficient quality of the evidence. Included in the review were interventions containing ABA practices, such a milieu communication therapy. Only one review focused on ABA approaches to intervention specifically. In a narrative review, Feeley and Jones (2008) outline the benefits of ABA for addressing challenging behavior in

children with Down syndrome. These previous reviews suggest that there is a sufficient number of behavioral intervention studies to support this approach to intervention for DS and a synthesis of the available evidence on behavioral approaches to intervention for individuals with DS is needed.

Although these reviews note support for ABA interventions, the limitations warrant an updated review and meta-analysis of ABA intervention effectiveness for individuals with DS. One of the previous reviews was narrative (Feeley & Jones, 2008) and thus was unsystematic in methodology and presentation of results. The other previous reviews did not focus on ABA approaches to intervention specifically (NYSDOH, 2006; Neil & Jones; 2018) limiting the ability to further analyze characteristics of the intervention. Further, an updated rating of methodological characteristics is needed. Previous systematic reviews either did not classify the quality of the evidence (Neil & Jones, 2018) or are dated (NYSDOH, 2006). Finally, only one of the previous reviews calculated effect sizes (Neil & Jones, 2018) using percentage non-overlapping data which has well-documented limitations (Kratochwill et al., 2010; Parker, Vannest, & Davis; 2011). A systematic review and meta-analysis that addresses these limitations is needed to provide a comprehensive evaluation of the ABA literature applied to the needs of individuals with DS.

Single-Case Research Design and Quality

Historically, the most common research designs employed in the field of ABA have been single-case designs (Kennedy, 2005) and, therefore, is especially relevant to reviews of ABA. While there is a long history of tools to evaluate the quality of group design research, only recently have experts begun to develop quality indicators and standards which can be used to synthesize the methodological characteristics of single-case design research. The What Works

Clearinghouse (WWC; Kratochwill et al., 2010) standards and the Risk of Bias in *N*-of-1 Trials scale (RoBiNT; Tate et al., 2013) are two such tools developed to characterize the quality of single-case designs.

In 2010, WWC assembled a panel of experts in single-case design and analysis to develop research quality standards for single-case designs (Kratochwill et al., 2010). The RoBiNT (Tate et al., 2013) is a 15-item instrument used to measure risk of bias in single-case reports and is designed to align with Kratochwill and colleagues (2010) published standards for single-case research. Since WWC and the RoBiNT place emphasis on well-defined objective features, they have the potential to increase the reliability and transparency of the synthesis of the results. Furthermore, WWC is also one of the only instruments that uses the same rating system for group designs, allowing researchers to make reasonable comparisons across studies (Kratochwill et al., 2010). In the current review, we extend previous reviews of intervention for individuals with DS by applying these quality indicators to the literature.

Effect Sizes

Quantitative syntheses are an important part of establishing the evidence-base for effective behavioral interventions and the systematic analysis of functional relations between independent and dependent variables (Kratochwill et al., 2013; Parker & Hagan-Burke, 2007). Nonoverlap effect sizes are non-parametric statistical methods used to analyze single-case research designs and are suited for the characteristics of single-case research studies compared to parametric measures (Parker et al., 2011b). These statistics measure the extent of nonoverlap of the data between adjacent phases (Parker et al., 2011a). Tau-U is an effect size that measures nonoverlap in adjacent phases while also correcting for trend (Parker et al., 2011a, b). Compared to other nonoverlap approaches and regression, Tau-U can account for the level of intervention

effectiveness with greater statistical power (Parker et al, 2011b). Tau-U has recently been included in a number of meta-analyses of single-case research to assess study outcomes (e.g., Ninci et al., 2019).

Purpose

The current study addresses limitations in previous reviews and expands on the current literature by (a) systematically reviewing ABA approaches to intervention applied to the behavior of individuals with DS; (b)evaluating the quality of single-case design studies using WWC and the RoBiNT; and (b) calculating Tau-U effect sizes to measure efficacy of single-case interventions meeting WWC design standards (Kratochwill et al., 2010, 2013) This review and meta-analysis addresses the following research questions:

1. What are the characteristics of ABA interventions (i.e. components, target skill domain) for individuals with DS?
2. What is the research design quality of studies of the effectiveness of ABA interventions for individuals with DS?
3. What is the magnitude of the effect of ABA interventions meeting quality standards for individuals with DS?

Method

The review protocol was registered online with the International Prospective Register of Systematic Reviews (Registration number: CRD42018089688). Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) standards are followed in this report (See Appendix A).

Study Identification

Search Strategy. The following databases were searched for peer-reviewed studies: ERIC, PsycINFO, MEDLINE, PubMed. Search terms for applied behavior analysis were determined using the Pearl Harvesting Information Retrieval Framework (Sandieson, 2006; Sandieson et al., 2010). Table 1 shows the search terms. The search was restricted to articles published up to November 2019, the date when the last searches were run. Studies were also identified by a hand search of journals which commonly publish single-case research with individuals with developmental disabilities in February 2018: *Research in Developmental Disabilities, Augmentative and Alternative Communication, Journal of Applied Behavior Analysis, Down Syndrome Research and Practice, and Behavior Modification*. Reference lists of included articles were searched to find additional relevant articles.

Inclusion Criteria. The following inclusion criteria were applied: (a) publication in English in a peer-reviewed journal, (b) original, empirical data on an ABA treatment; (c) involving at least one participant with DS without a comorbid diagnosis of autism spectrum disorder (where participants had a range of differing diagnoses, the study was included if at least one participant met the inclusion criteria and outcome data were presented separately for the participant with DS); (d) employed a single-case experimental design; and (e) included a linear graphical presentation of data.

Articles were included if they explicitly stated the use of one or more of the behavior analytic components listed by Wong et al. (2015). During review, several additional strategies were added after reaching consensus between coders. The final list of included strategies is presented in Table 1. Articles were excluded if they focused solely on the outcomes of an assessment (e.g. functional assessment, preference assessment) without a corresponding intervention plan.

Title/Abstract and Full Text Review. The flow diagram of study selection is presented in Figure 1. All references were exported excel where any duplicates were removed. Initially, two authors (NN, BA) independently screened a random sample of 66 titles and abstracts in which they were blinded to authors and journal titles and reached strong agreement (Cohen's k = 0.81). They then independently screened all titles and abstracts and disagreements were discussed until consensus. Where there was doubt whether the article met inclusion criteria from the title and abstract alone, it was included for full-text review. The full text of all articles considered for inclusion were independently reviewed by the same two authors (NN, BA) and consensus was reached via discussion.

Variable Coding

Information on pre-specified study components including participant and setting characteristics, the dependent variable, intervention components, and the design and methodological quality was extracted using a form created for this study.

Participant and Setting Characteristics. Authors recorded the total number of participants, the number of participants with DS and for each participant with DS, the age, gender, IQ, and any co-occurring diagnoses/conditions. The reported setting was also recorded.

Dependent Variable. Items in this category included (a) the primary dependent variable, (b) the skill domain, and (c) the direction of expected change. The skill domain was categorized according to the operational definitions presented in the Ontario Scientific Expert Taskforce for the Treatment of Autism Spectrum Disorder Report (ONTABA, 2017), listed in Appendix B.

Intervention Components. Authors coded intervention components as present or not present, using definitions from Wong et al. (2015). Authors noted any additional behavior analytic components of interventions via a write in "Other" code. An additional three codes were

created from this: Punishment/overcorrection, time out, and posture Training. They also coded the role of the individual implementing intervention as researcher, teacher, parents/primary caregiver, peer, or paraprofessional.

Design and Methodological Quality. Scientific quality of the studies was evaluated using the RoBiNT scale (Tate et al., 2013) for single-case experiments. The RoBiNT scale is a 15-item rating scale designed to facilitate the critical appraisal of single-case intervention studies. It is a revision of the Single-Case Experimental Design (SCED) scale, developed by the same authors (Tate et al., 2008). The RoBiNT has two subscales: internal validity and external validity and interpretation. The RoBiNT scale is a reliable and valid measure (Tate et al., 2013). A score of two is given to studies that meet the revised stringent standards representing those recommended by authorities in the field. A score of one is given to studies that meet the original SCED criteria or are in a position between completely satisfying and not satisfying the criterion at all (many of these also correspond to "meeting standards with reservations" on the WWC standards). A score of zero is given to studies that do not meet the revised or original SCED criteria (Tate et al., 2013). Specifically, study of the inter-rater reliability with experienced raters using the scale yielded respectable intra-class correlation coefficients for the total score and both subscales (total score = .90, internal validity = .88, and external validity and interpretation = .87). Similar results were obtained when trained novice raters used the scale and there is research to support the RoBiNT's construct validity (Tate et al., 2015).

Application of WWC Evidence Standards. Authors recorded whether each study met methodological criteria described in the WWC standards (2017): (a) the independent variable was systematically manipulated, (b) three attempts to demonstrate treatment effects at three different points in time; (c) reliability data collected, measured for at least 20% of sessions

overall and in each condition, and met minimal acceptability standards of 80% agreement; (d) data point sufficiency. Studies were rated as "Meets Standards" if five or more data points were present in each phase, as "Meets Standards with Reservations" if three to four data points were present in each phase, and "Does Not Meet Standards" if there were fewer than three data points per phase. Multiple probe designs had to meet additional criteria: (a) initial baseline sessions must overlap vertically (3 consecutive probe points for each case to Meet WWC Single-Case Design Standards without Reservations or at least 1 probe point for each case to Meet WWC Single-Case Design Standards with Reservations); (b) Probe points must be available just before introducing the independent variable (Within the 3 sessions just before introducing the independent variable, the design must include 3 consecutive probe points for each case to Meet WWC Single-Case Design Standards without Reservations or at least 1 probe point for each case to Meet WWC Single-Case Design Standards with Reservations) and; (c) each case not receiving the intervention needs a probe point in a session in which another case either first receives the intervention or reaches the prespecified intervention criterion. In multiple baseline across participants designs where only one or two participants had a diagnosis of DS, consideration for all participants was used to assess the overall methodological standards of the study although the participants without DS were excluded from the analysis of participant characteristics and data extraction.

Data Extraction and Effect Size Calculation

Data Extraction. For all cases meeting design standards with or without reservations, data were extracted from each graph using WebPlotDigitizer (Rohatgi, 2019), a free browser-based application for extracting and exporting raw data from different types of plots. Only participants with the diagnosis of DS were included in effect size calculation. Individuals with

comorbid diagnosis of DS and other developmental disabilities or individuals without a diagnosis of DS were included in the quantitative synthesis. The raw data were entered into separate columns in a spreadsheet by phase. Extracted data were recoded so that positive effect sizes represented behavioral changes in the desired direction.

Data Analysis. Tau –U was used to quantify change as a result of the intervention (Parker, Vannest, & Davis, 2011; Parker, Vannest, Davis, & Sauber, 2011). Tau-U is the percentage of data showing improvement across phases calculated by comparing pairs of data points. First, we calculated Tau-U for each AB phase contrast (e.g., an effect size for the A1/B1 contrast and a separate effect size for the A2/B2 contrast) for each dependent variable using a free, online Tau-U calculator (Vannest, Parker, & Gonan, 2011) and then combined these experiment-level effect sizes into one omnibus effect size per study. Baseline data were denoted as "Phase A" and intervention data (even if intervention occurred over multiple phases) were denoted as "Phase B." Maintenance and generalization data were not included in "Phase B". For each graph, the calculation of the comparison between phases controlled for positive baseline trend to provide a conservative estimate of the "true effect". When combining Tau-U values, an inverse weighting scheme was used that gives more credit to studies with more data points. Currently, the Tau-U metric supports inclusion of multiple baseline and reversal designs; however, their use for other single-case experimental design such as alternating treatment designs and changing criterion designs does not have a sound research base (Shadish, Hedges, Horner, & Odom, 2015). Parker et al. (2011) provides the following interpretation guidelines for Tau-U: small effect = 0 to 0.62; medium effect = 0.63 to 0.92; large effect = 0.93 to 1.00.

Inter-rater Reliability

Variables for each included study was initially coded by the first author. To assess

interrater reliability of the variable coding, 30.4% of the included articles were independently coded by the second author. She coded information independently using the same form as the first coder. Initial agreement ranged from 94-100% for participant, setting, dependent variable and intervention coding, 76-100% for RoBiNT coding, and 84% for WWC evaluation. If there were disagreements, they were discussed until they came to a consensus, resulting 100% final agreement across all codes.

For data extraction, an undergraduate research assistant extracted the graphs from 25% of the articles included in the meta-analysis using Webplot digitizer. Disagreements were counted when the whole number did not match, decimal places were not considered. For data extraction scores that were counted as a disagreement, the files and graphs were reviewed to determine which was accurate. Agreement was 89.6%; there was only one set of data with disagreements, which resulted from improper calibration of the y-axis.

For data analysis, an undergraduate research assistant entered the extracted AB contrast data for 25% of the articles included in the meta-analysis into the Tau-U calculator and effect sizes obtained for the first and second raters were compared for reliability purposes, agreement was 100%.

For all steps, interrater reliability was calculated by dividing the number of agreements by the number of agreements plus disagreements and multiply by 100 to obtain a percentage.

Results

A total of 127 studies were identified for inclusion in the review. Studies were published from 1997 to 2019. See Appendix C for a list of included documents. An overview of the search process with reasons for removing ineligible studies is presented in Figure 1.

Participant Characteristics

Studies included 201 participants. Eighty-three (41.3 %) participants were female, 111 (55.2 %) were male, and for 6 (3.0 %) the participants gender was not reported. The mean age was 12.98 (SD = 11.58) ranging from 5 months to 54 years old. Participants were evenly distributed across age groups; 55 participants were in early childhood (< 5 years), 60 were in middle childhood (5-11), 40 were in adolescence (12-18) and 46 were chronological adults (> 18). An IQ score was reported for 67 (33.3 %) of the participants; the mean IQ score was 46.67 (SD =14.59). Of the 201 participants examined in this study, 74 had one or more comorbid diagnoses reported, including: intellectual disability (n = 53), developmental delay (n = 5), auditory impairment (n = 6), visual impairment (n = 2), speech/language impairment (n = 3), epilepsy/seizure disorder (n = 3), sensory impairment (n = 6), physical disability (n = 4), stereotypic movement disorder (n = 1), and gastro-esophageal disorder (n = 1).

Setting and Implementers

Studies took place in a variety of settings, including school (n = 53), home (n = 21), clinic (n = 34), community (n = 8), group home or day program (n = 17), Clinic or University (n = 33. In seven of the studies, the setting was not specified. The implementers included: researchers (n = 79), teachers (n = 26), parents/primary caregivers (n = 14), peers (n = 3), paraprofessionals (n = 10), and one study used a technological aided intervention (n = 1).

Dependent Variables

The most commonly targeted skill domains were communication (n = 37) and challenging behavior (n = 32) followed by personal responsibility/adaptive (n = 12), motor (n = 8), social/interpersonal (n = 5), learning/school readiness (n =6), academic (n = 4), play (n = 2), vocational (n = 2). Nineteen studies investigated more than one skill domain.

In the studies examined, various methods were used to collect data on the dependent variables—44.8% of studies used trial-based measures (n = 57), 29.13% used frequency/rate-based measures (n = 37), 22.8% used time sampling (n = 29), 6.3% used duration (n = 8), and 2.4% used permanent product (n = 3). Thirty-one studies targeted decreasing behavior, 77 studies increasing behavior, and 18 studies increasing and decreasing.

Intervention Components

As shown in Table 2, most studies included approximately 3 intervention components. Studies with outcomes in communication and mixed areas had the greatest number of intervention components per study (3.39 and 3.5) where social/interpersonal and motor had the fewest per study (2.4 and 2.71). For all skill domains, the most common intervention component was reinforcement (105). Prompting the second most frequent intervention component for all skill domains except challenging behavior, where differential reinforcement was the second most common component. The third most common intervention component varied across skill domains. For communication and social/interpersonal it was discrete trial teaching/training, for personal responsibility/adaptive and motor it was task analysis, for academic, play, and vocational domains it was modeling. In the learning and school readiness and challenging behavior domains it was antecedent based intervention and for mixed domains, it was FBA. Several studies used restrictive approaches including punishment (8) and restraint fading (2).

Across age groups, reinforcement and prompting and were the top two most commonly used intervention components. For chronological adults (>18), DRA/I/O was the third most commonly included component. For adolescents (12-18), antecedent-based intervention was the third most commonly included component and for middle (5-11) and early childhood (<5), DTT was the third most common component.

RoBiNT Scale

The number of studies receiving ratings of 0, 1, and 2 on the internal validity and external validity subscales of the RoBiNT (Tate et al., 2013, 2016) are presented in Figure 2. Regarding the internal validity of the included studies, relative strengths were found in selecting designs which have three opportunities to observe experimental effects, sampling at least three data points in each phase, and the inclusion of interrater agreement. Few studies used blinding of participants (n = 1) or assessors (n = 3) or included a measure of treatment adherence (n = 33). On the external validity subscale, few studies included measures of generalization (n = 22), a replication of experimental procedures with additional participants with DS (n = 24), or systematic visual analysis (n = 3). Most studies provided data for all sessions (n = 116), and intervention and measurement procedures were generally well described.

WWC Design Standards

Only 21 studies met the WWC standards without reservation, 25 met with reservation, and the remaining 81 did not meet the standards (See Table 3). Of the studies aimed at increasing behavior (n = 77), 15 met with reservation, 11 met without reservation, and 51 did not meet. Of those aimed at decreasing behavior (n = 32), 6 met with reservation, 7 met without reservation, and 19 did not meet. Finally, of the studies targeting both (n = 18), 3 met with reservation, 3 met without reservation, and 11 did not meet.

Quantitative Synthesis

Thirty-seven studies were included in the meta-analysis, nine were excluded due to inappropriate designs (i.e. alternating treatment). A total of 56 participants produced 52 experiments and 156 AB contrasts between phases A and B for the calculation of effects. Tau-U effect sizes for ranged from 0.25 to 1.00. The overall effect of ABA intervention on behavior of

children with DS was 0.89 (90 % CI = [0.85, 0.92]). See Table 4 for Tau-U ESs by study and Figure 3 for a forest plot. An effect of this size and CI indicate medium effects of the intervention on dependent variable outcomes and may be interpreted as 89% of the intervention phase data are improved above baseline levels (corrected for undesirable trend).

Conclusion

This study reviewed and meta-analyzed studies using ABA techniques to increase and decrease behavior in individuals with DS with the following aims: to identify the characteristics of ABA interventions (i.e. components, target skill domain) for individuals with DS, summarize the research design quality of studies of the effectiveness of ABA interventions for individuals with DS and identify the magnitude of the effect of ABA interventions for individuals with DS meeting WWC standards. One-hundred twenty-five studies were included and evaluated for their quality. Thirty-seven studies meeting WWC standards were identified and demonstrated a medium overall effect size. These results suggest the use of ABA-based interventions are promising for behavior change in individuals with DS, but more work is needed to establish effects using high quality research for a broad range of outcomes.

We first sought to characterize the types of interventions applied to the behavior of individuals with DS. Fundamental ABA techniques made up the majority of commonly used intervention components: reinforcement, and prompting (Cooper, Heron, & Heward, 2007). Reinforcement refers to a stimulus (activity, event) occurring after a learner engages in a behavior resulting in increased occurrence of the behavior in the future. Prompting is assistance (verbal, gestural, physical) provided to a learner to assist them in engaging in a targeted behavior. Prompts are generally given by an adult or peer before or as a learner attempts to use a skill.

Given the majority of the studies included in this analysis involved increasing behavior, it is understandable that studies rely on these fundamental techniques to teach these new skills.

Most of these components were delivered as an intervention package to address a specific target behavior for a participant. For example, Hildebrand et al (1990) used physical and visual prompts, modeling, self-recruited reinforcement, a task analysis, and self-management to teach a 36-year-old male with DS to complete a vocational task (folding paper). Only one study, Poulson (1988) used reinforcement alone to increase the vocalizations of infants with DS. These fundamental techniques have strong support for their use across various populations and are considered evidence-based practice among other developmental disabilities (Wong et al., 2015), however, continued research will be needed to determine the effects of various packages made up of these techniques.

Intrusive procedures, such as punishment and restraint fading were also present within the literature. Generally, as researchers and practitioners have begun to develop and rely on more positively oriented interventions and functional assessment, intrusive procedures have tended to be used and reported less frequently in the literature (Matson & LoVullo, 2008). This trend is also consistent within our database, with the number of studies reporting intrusive procedures decreasing over time.

Although there were many studies including individuals with DS, less than half of the studies focused on individuals with DS and their specific needs. For most of the studies, individuals with DS were included as part of a broader population of individuals with intellectual or developmental disabilities and represented one participant among a heterogenous sample. Studies with heterogenous groups also tended to illustrate the utility of a particular intervention component or package, without regard to the diagnostic characteristics of the participants. Thus,

a broad range of outcomes relevant to individuals with intellectual and developmental disabilities were targeted within the studies and were not specific to needs of individuals with DS. For example, a large proportion focused on communication. Language development, and particularly expressive language development is delayed and often shows persistent problems into adulthood for individuals with DS (Fidler 2005) but is also a persistent problem for individuals with other developmental and intellectual disabilities. With this knowledge, it may seem reasonable to target these skills for a heterogeneous group of individuals with intellectual disability. However, the language profile of individuals with DS differs from their nonverbal mental age, suggesting that the language difficulties are not a result of their intellectual disabilities and may be more similar to the pattern seen in children with specific language impairment (Naess et al. 2011). Since their language profile differs, there may be a benefit to interventions designed to meet the specific needs of individuals with Down syndrome.

Applying interventions to a broad range of individuals can have limitations. Individuals with genetic syndromes, such as DS, display increased probability of displaying associated behaviors, termed the behavioral phenotype (Fidler, 2005). Several researchers have highlighted how behavioral phenotypes can be used to guide the development of interventions targeting the critical needs of individuals with DS and may result in improved outcomes (e.g. Dykens, 1995, Fidler, 2005). Behavioral researchers have conceptualized the behavioral phenotype as a biological setting event that functions as a motivating operation to change the value of certain consequences and behavior associated with those consequences (Bauer & Jones, 2014). Feeley and Jones (2006) suggest that DS itself functions as establishing operation, increasing the aversiveness of tasks and the value of escape as a reinforcer and increasing the likelihood of behavior to escape tasks compared with typically developing individuals. Further, individuals

with DS have a higher propensity for factors commonly considered to be setting events for challenging behavior, such as an increased likelihood for sleep issues (Stores, 1993) or high incidence of illness in children with DS (Roizen, 1996).

Recent work on reading interventions uses the behavioral phenotype literature to tailor intervention to the DS behavioral phenotype with positive results (e.g. Lemons et al., 2017). In their intervention, they capitalize of the visual processing strengths reported in DS by including highly imageable sight words and pairing them with pictures. Cognitive weaknesses were addressed by minimizing complexity of instruction and embedding scaffolding. Language weaknesses were addressed by allowing nonverbal responses. This tailored approach was notably absent from the literature in this review. A simple addition would be additional visual support to draw on the relative visual processing strengths of individuals with DS (as in Lemons et al., 2017) and yet only four studies reported their inclusion. Combining the phenotype literature in the design of future ABA-based interventions may result in more effective approaches for individuals with DS and future research should pursue this line of work. For example, if it is known that DS functions to increase the aversiveness of tasks and task-related consequences, interventionists might decrease the difficulty of tasks via task analyses and prompting and pair highly valued social interaction for engaging in those tasks which, after repeated pairings, should make task-related consequences function as reinforcers.

Similarly, a number of studies included in the review involved adolescent and adult participants, yet there was a lack of emphasis on skills specific to individuals with DS as they age. Few studies targeted phenotypic needs of adolescents and adults with DS (Chapman & Heskseth, 2000) such as grammar and syntax, and no studies focused on outcomes related to dementia such as self-care skills, mental health concerns, or behavioral excesses. There is a need

for treatment research targeting the needs adult and adolescents with DS as individuals with DS continue to need support in living independently and working in competitive employment settings (Esbensen, Bishop, Seltzer, Greenberg, & Taylor, 2010).

The second aim of this review was to evaluate the of quality single-case design studies applying ABA to the behavior of individuals with DS. Both the quality measures used in this review highlighted methodological weaknesses in the literature. Fewer than half the designs were met WWC criteria with or without reservations. Future single-case research should rely on strong experimental design which are more likely to demonstrate experimental control, ensuring to have three possible demonstrations of intervention effect. While studies frequently reported observer agreement, they often failed to collect a sufficient proportion during all phases.

Studies often did not meet WWC standards because of sampling of behavior. Meeting with reservations requires three data points in each phase, compared to five to meet without reservations. Too few data points can mean a data series too short to examine level, trend, variability, and related properties of the data, threatening data interpretation (Kratochwill, 2010). Even still, how many data points are needed to establish stability is dependent on multiple features of the data series including variability, expected intervention effects, the target response, and method of measurement.

The WWC does not include several of the indicators of the quality included on the RoBiNT. First, it does not include a measure of fidelity as a quality indicator. When studies do not report treatment fidelity it is not known whether the treatment was carried out as planned raising doubt as to whether the independent variable affected the changes in the dependent variable. Some quality indicators consider treatment fidelity to be a key indicator (e.g. Horner et al., 2005), while others do not (e.g. Reichow et al. 2008). Analysis with RoBiNT revealed the

majority of studies did not include an adequate measure of fidelity. Attending to fidelity in future research will help to determine the effects of ABA in addressing behavior for children with DS with more confidence.

Other areas captured by the RoBiNT but not by the WWC are several indicators of external validity. The heterogenous nature of the samples within studies resulted in lowered scores on the external validity scales of the RoBiNT. Studies often failed to replicate procedures across multiple individuals with DS. Replication decreases the likelihood that a particular extraneous variable will affect outcomes and increase generalizability. Further research will be needed in order to replicate the efficacy of included procedures for this population. Almost no studies used a systematic means of conducting visual analysis, and small number included measures of generalization. One standard of external validity absent from the RoBiNT is the measurement of social validity. Researchers should also consider the social validity, or importance of goals, outcomes and procedures to their participants.

Third, we sought to identify the overall effect of studies meeting WWC standards using single-case meta-analysis. The overall Tau-U was medium from the studies included in the quantitative synthesis, although Tau-U varied widely between experiments. These results provide preliminary evidence that ABA interventions are effective at addressing the behavior of individuals with DS. While the overall effect size was medium, it is primarily an effect size representing the outcomes of studies targeting challenging behavior and communication. No studies targeting academic skills, play, or vocational skills were eligible for inclusion in the quantitative syntheses. This points to the need for future research to conduct high quality studies investigating the effects of behavioral interventions targeting these areas other than communication and challenging behavior.

Readers should be aware that effect sizes alone do not and cannot replace visual analysis in single-case design research. Tau-U has been found to be consistent with visual analysis (Parker, Vannest & Davis, 2011) and has demonstrated convergent validity with visual analysis in studies (Ninci et al., 2015). Tau-U provides a measure of the magnitude of the effect of the studies included. It represents the improvement in the target behavior between baseline and intervention but does note evaluate causal relationships between the intervention and the target behavior. Effect sizes should always be considered alongside visual analysis.

Limitations

This review has several notable limitations. This study only included journals published in peer-reviewed journals, as we were interested in the design quality of published studies. In order to determine publication bias, there is a need to evaluate "grey literature" (e.g., unpublished studies). Nonetheless, we feel our emphasis on published single-case design studies meeting methodological quality standards represents a unique contribution to the literature. Second, because many behavioral components often made up a package intervention, it is likely that some studies were missed because it was not clear that ABA practices were a primary component of the intervention. We used a broad range of search terms to capture as many studies as possible but relied on authors explicit identification of approaches when determining inclusion. Third, data on generalization and maintenance conditions were not analyzed and it is unclear whether interventions would result in meaningful changes outside of the treatment context or over extended periods of time. Fourth, this study did not analyze experiments via visual analysis. Tau-U is useful for combining the results of a body of literature, yet there can be concerns when using it in isolation, and readers should not equate a large effect with demonstration of experimental control (Brossart, Vannest, Davis, & Patience, 2014).

Additionally, as with any nonparametric measure of effect, Tau-U has some limitations including ceiling effects (Parker et al. 2011a, b) with several included studies hitting the upper ceiling. Finally, although we only included studies which met WWC standards with or without reservations, results should be interpreted in light of the design weaknesses. Few studies included a measure of procedural integrity which threatens the internal validity of even the included studies.

Future Research

Future research should investigate several areas. Additional research is needed which is tailored to the specific needs of individuals with DS. Research is also needed which investigates interventions for the needs of adolescents and adults with DS; in particular, the needs associated with dementia. Greater emphasis on generalization and maintenance conditions are needed in order to determine whether acquired skills are generalized and maintained across settings, people and time. Future research should also include component analyzes of interventions in order to determine the additive effects of various techniques.

Future research should also address the methodological weaknesses of the current literature and address the standards of quality outlined by the RoBiNT (Tate et al., 2013) and WWC (Kratochwill et al., 2010), as well as other standards such as the Council for Exceptional Children (Cook et al., 2015). Future studies should assess convergence of visual analysis and effect sizes, such as Tau-U, and may also consider using more than one effect size.

References

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

Brossart, D. F., Vannest, K. J., Davis, J. L., & Patience, M. A. (2014). Incorporating nonoverlap indices with visual analysis for quantifying intervention effectiveness in single-case experimental designs. *Neuropsychological Rehabilitation, 24*(3-4), 464-491. doi: 10.1080/09602011.2013.868361

Buckley, S., Bird, G., Sacks, B., & Archer, T. (2006). A comparison of mainstream and special education for teenagers with Down syndrome: Implications for parents and teachers. *Down Syndrome Research and Practice, 9*(3), 54-67.

Cadogan, S., & McCrimmon, A. W. (2015). Pivotal response treatment for children with autism spectrum disorder: A systematic review of research quality. *Developmental Neurorehabilitation, 18*(2), 137-144. doi:10.3109/17518423.2013.845615

Chapman, R. S., & Hesketh, L. J. (2000). Behavioral phenotype of individuals with Down syndrome. *Mental Retardation and Developmental Disabilities Research Reviews*, *6*(2), 84-95. doi: 10.1002/1098-2779(2000)6:2<84::AID-MRDD2>3.0.CO;2-P

Cook, B. G., Buysse, V., Klingner, J., Landrum, T. J., McWilliam, R. A., Tankersley, M., & Test, D.W. (2015). CEC's standards for classifying the evidence base of practices in special education. *Remedial and Special Education, 36*(4), 220–234. doi: 10.1177/0741932514557271.

Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). *Applied Behavior Analysis* (2nd edn.). Upper Saddle River, NJ: Pearson.

Crook, N., Adams, M., Shorten, N., & Langdon, P. E. (2016). Does the well being of individuals with Down Syndrome and dementia improve when using life story books and rummage boxes? A randomized single case series experiment. *Journal of Applied Research in Intellectual Disabilities, 29*(1), 1-10. doi: 10.1111/jar.12151

Dykens, E. M. (1995). Measuring behavioral phenotypes: Provocations from the" new genetics.". *American Journal on Mental Retardation, 99*(5), 522-532.

Esbensen, A. J., Bishop, S., Seltzer, M. M., Greenberg, J. S., & Taylor, J. L. (2010). Comparisons between individuals with autism spectrum disorders and individuals with Down syndrome in adulthood. *American Journal on Intellectual and Developmental Disabilities*, *115*(4), 277-290. doi: 10.1352/1944-7558-115.4.277

Fidler, D. J. (2005). The emerging Down syndrome behavioral phenotype in early childhood: Implications for practice. *Infants and Young Children*, *18*(2), 86-103. doi: 10.1097/00001163-200504000-00003

Gottfredson, D. C., Cook, T. D., Gardner, F. E. M., Gorman-Smith, D., Howe, G. W., Sandler, I. N., & Zafft, K. M. (2015). Standards of evidence for efficacy, effectiveness, and scale-up research in prevention science: Next generation. *Prevention Science, 16*(7), 893–926. doi: 10.1007/s11121-015-0555-x.

Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. Exceptional children, 71(2), 165-179.

Jacola, L. M., Hickey, F., Howe, S. R., Esbensen, A., & Shear, P. K. (2014). Behavior and adaptive functioning in adolescents with Down syndrome: specifying targets for

intervention. *Journal of Mental Health Research in Intellectual Disabilities, 7*(4), 287-305. doi: 10.1080/19315864.2014.920941

Johnston, M. V., & Smith, R. O. (2010). Single subject designs: Current methodologies and future directions. *OTJR: Occupation, Participation, Health*, *30*(1), 4-10. doi: 10.3928/15394492-20091214-02.

Kasari, C. (2015). Update on behavioral interventions for autism and developmental disabilities *Current Opinion in Neurology*, *28(2)*, 124-129. doi:10.1097/WCO. 0000000000000185

Kennedy, C. H. (2005). *Single-case designs for educational research*. Boston, MA: Allyn & Bacon.

Kennedy, E.J., & Flynn, M.C. (2003). Training phonological awareness skills in children with Down syndrome. *Research in Developmental Disabilities, 24*, 44-57.

Kratochwill, T. R., Hitchcock, J. H., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2013). Single-case intervention research design standards. *Remedial and Special Education, 34*(1), 26-38.

Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2010). Single-case designs technical documentation. Retrieved from What Works Clearinghouse website: http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf.

Lemons, C. J., King, S. A., Davidson, K. A., Puranik, C. S., Al Otaiba, S., Fulmer, D., ... & Fidler, D. J. (2017). Developing an early reading intervention aligned with the Down syndrome behavioral phenotype. *Focus on Autism and Other Developmental Disabilities, 32*(3), 176-187. doi: 10.1177/1088357615618941

Lory, C., Rispoli, M., & Gregori, E. (2018). Play interventions involving children with autism spectrum disorder and typically developing peers: A review of research quality. *Review*

Journal of Autism and Developmental Disorders, *5*(1), 78-89. doi:10.1007/s40489-017-0124-2

Maggin, D. M., Briesch, A. M., & Chafouleas, S. M. (2012). An application of the what works clearinghouse standards for evaluating single-subject research: Synthesis of the self-management literature base. *Remedial and Special Education*, *34*(1), 44-58. doi:10.1177/0741932511435176

Martinez, J. R., Werch, B. L., & Conroy, M. A. (2016). School-based interventions targeting challenging behaviors exhibited by young children with autism spectrum disorder: A systematic literature review. *Education and Training in Autism and Developmental Disabilities*, *51*(3), 265–280.

Matson, J. L., & LoVullo, S. V. (2008). A review of behavioral treatments for self-injurious behaviors of persons with autism spectrum disorders. *Behavior Modification*, *32*(1), 61-76. doi: 10.1177/0145445507304581

Neil, N., & Jones, E. A. (2018). Communication intervention for individuals with Down syndrome: Systematic review and meta-analysis. *Developmental Neurorehabilitation, 21*(1), 1-12. doi: 10.1080/17518423.2016.1212947

New York State Department of Health (NSYDOH) (2006). Report of the Recommendations: Down Syndrome Assessment and Intervention for Young Children (age 0–3 Years). New York, NY: Author.

Ninci, J., Neely, L. C., Burke, M. D., Rispoli, M., Vannest, K. J., Willson, V., & Ulloa, P. (2019). A meta-analysis of single-case research on embedding interests of individuals with autism spectrum disorder. *Remedial and Special Education*. doi: 10.1177/0741932519841692

Ninci, J., Neely, L. C., Hong, E. R., Boles, M. B., Gilliland, W. D., Ganz, J. B., ... & Vannest, K. J. (2015). Meta-analysis of single-case research on teaching functional living skills to individuals with ASD. Review Journal of Autism and Developmental Disorders, 2(2), 184-198.

ONTABA Scientific Expert Task Force for Treatment of Autism Spectrum Disorders (OSETT-ASD) (2017). Evidence-based practices for individuals with autism spectrum disorder: Recommendations for caregivers, practitioners, and policy makers. Toronto, ON. Retrieved from http://www.ontaba.org/pdf/ONTABA%20OSETT-ASD%20REPORT%20WEB.pdf

Parker, R. I., & Hagan-Burke, S. (2007). Useful effect size interpretations for single case research. *Behavior Therapy, 38*(1), 95-105. doi: 10.1016/j.beth.2006.05.002

Parker, R. I., Vannest, K. J., & Davis, J. L. (2011). Effect size in single-case research: A review of nine nonoverlap techniques. *Behavior Modification, 35*(4), 303-322. doi: 10.1177/0145445511399147

Parker, R. I., Vannest, K. J., Davis, J. L., & Sauber, S. B. (2011). Combining nonoverlap and trend for single-case research: Tau-U. *Behavior Therapy, 42*(2), 284-299. doi: 10.1016/j.beth.2010.08.006

Reichow, B., Volkmar, F. R., & Cicchetti, D. V. (2008). Development of the evaluative method for evaluating and determining evidence-based practices in autism. Journal of autism and developmental disorders, 38(7), 1311-1319.

Rohatgi, A. (2019). WebPlotDigitizer (Version 4.2) [Computer software]. Retrieved from https://automeris.io/WebPlotDigitizer/

Roizen, N. J. (1996). Down syndrome and associated medical disorders. *Mental Retardation and Developmental Disability Research Reviews,* 2*,* 85-89.

Shadish, W. R., Hedges, L. V., Horner, R. H., & Odom, S. L. (2015). The role of between-case effect size in conducting, interpreting, and summarizing single-case research. Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education.

Stores, R. (1993). A preliminary study of sleep disorders and daytime behavior problems in children with Down's syndrome. *Down Syndrome Research and Practice,* 1, 29-33.

Tate, R. L., McDonald, S., Perdices, M., Togher, L., Schultz, R., & Savage, S. (2008). Rating the methodological quality of single-subject designs and n-of-1 trials: Introducing the Single Case Experimental Design (SCED) scale. *Neuropsychological Rehabilitation, 18*(4), 385–401. doi: 10.1080/09602010802009201

Tate, R. L., Perdices, M., Rosenkoetter, U., Wakim, D., Godbee, K., Togher, L., & McDonald, S. (2013). Revision of a method quality rating scale for single-case experimental designs and n-of-1 trials: The 15-item Risk of Bias in N-of-1 Trials (RoBiNT) Scale. *Neuropsychological Rehabilitation*, *23*(5), 619, doi:10.1080/09602011.2013.824383

Tate, R., Rosenkoetter, U., Wakim, D., Sigmundsdottir, L., Doubleday, J., Togher, L., McDonald, S., & Perdices, M. (2015). *The Risk-of-Bias in N-of-1 Trials (RoBiNT) Scale: An expanded manual for the critical appraisal of single-case reports*. Sydney: John Walsh Centre for Rehabilitation Research.

Ontario Scientific Expert Task Force for the Treatment of Autism Spectrum Disorder (OSETT-ASD) (2017). *Evidence-based practices for individuals with autism spectrum disorder:*

Recommendations for caregivers, practitioners, and policy makers. Ontario, Canada: The Ontario Association for Behavior Analysis (ONTABA).

Ulrich, D. A., Ulrich, B. D., Angulo-Kinzler, R. M., & Yun, J. (2001). Treadmill training of infants with Down syndrome: Evidence-based developmental outcomes. *Pediatrics, 108*, 42– 48. doi: 10.1542/peds.108.5.e84

Vannest, K. J., Parker, R. I., & Gonan, O. (2011). Single case research: Web based calculators for SCR analysis (Version 1.0) [Webbased application]. College Station: Texas A&M University. www.singlecaseresearch.org

Wong, C., Odom, S. L., Hume, K. A., Cox, A.W., Fettig, A., Kucharczyk, S., . . . Schult, T. R. (2015). Evidence-based practices for children, youth, and young adults with autism spectrum disorder: A comprehensive review. *Journal of Autism and Developmental Disorders, 45*(7), 1951–1966. doi: 10.1007/s10803-014-2351-z.

World Health Organization. (2017). Genes and human disease: Genes and chromosomal disease: Down syndrome. United States of America: World Health Organization. Retrieved from http://www.who.int/genomics/public/geneticdiseases/en/index1.html

Compliance with Ethical Standards

The manuscript does not contain clinical studies or patient data.

Tables

Table 1.

Search Terms for Study Identification

Database	Category	**Search Terms**
ERIC, PsychINFO, MEDLINE, PubMed	Down Syndrome	“Down* syndrome” OR “Trisomy 21”
		AND
	Applied Behavior Analysis (ABA)	"applied behavior* analysis" OR "functional behavior* assessment*" OR "token economy*" OR "abolishing operation*" OR "ABC checklist" OR "differential reinforce*" OR "abstinence reinforce*" OR "interobserver agreement" OR "schedule* of reinforce*" OR "backward chaining" OR "forward chaining" OR "different* reinforce*" OR "differential reinforce* of alternative behavior" OR "prompting behavior" OR "prompting behavior" OR "continuous reinforce*" OR "direct assessment*" OR "delay of reinforce*" OR "contingent escape" OR "functional analysis*" OR "function-based behavior*" OR "continuous reinforcement schedule*" OR "precision teaching" OR "antecedent intervention*" OR "function based intervention*" OR "discrete trial training" OR "extinction burst" OR "motivating operation*" OR "establishing operation*" OR "extinction schedule*" OR "fading prompt*" OR "functional alternative behavior*" OR "transition assessment*" OR "stimulus preference assessment*" OR "positive punishment*" OR "positive reinforce*" OR "negative punishment*" OR "negative reinforce*" OR “ABI” OR "applied behavioral intervention" OR “IBT” OR "intensive behavior intervention" OR “ABAI OR IBT” OR "discrete trial training" OR "behavior therapy" OR "early behavioral treatment" OR "functional behavioral analysis"

Table 2.

Intervention components by outcome variable type.

Component	COM	CB	PR/A	M	S/I	L/SR	A	P	V	Mix	Total
Number of Studies	37	32	12	8	5	6	4	2	2	1	127
Reinforcement (R+)	105	105	105	105	105	105	105	105	105	105	105
Prompting (PP)	79	79	79	79	79	79	79	79	79	79	79
Differential reinforcement of alternative/incompatible/other (DRA/I/O)	24	24	24	24	24	24	24	24	24	24	24
Discrete trial teaching/training	24	24	24	24	24	24	24	24	24	24	24
Antecedent-based intervention	19	19	19	19	19	19	19	19	19	19	19
Functional behavior assessment (FBA)	19	19	19	19	19	19	19	19	19	19	19
Modeling	16	16	16	16	16	16	16	16	16	16	16
Time delay	15	15	15	15	15	15	15	15	15	15	15
Extinction	12	12	12	12	12	12	12	12	12	12	12
Task analysis (TA)	9	9	9	9	9	9	9	9	9	9	9
Punishment procedures/overcorrection	8	8	8	8	8	8	8	8	8	8	8
Naturalistic intervention	7	7	7	7	7	7	7	7	7	7	7
Technology-aided instruction and intervention (TAII)	6	6	6	6	6	6	6	6	6	6	6
Parent-implemented intervention	5	5	5	5	5	5	5	5	5	5	5
Self-management (SM)	5	5	5	5	5	5	5	5	5	5	5
Visual support (VS)	4	4	4	4	4	4	4	4	4	4	4
Functional communication training (FCT)	6	6	6	6	6	6	6	6	6	6	6
Peer-Mediated Instruction	5	5	5	5	5	5	5	5	5	5	5
Video modeling (VM)	4	4	4	4	4	4	4	4	4	4	4
Response Blocking	3	3	3	3	3	3	3	3	3	3	3
Removal of Restraints	2	2	2	2	2	2	2	2	2	2	2
Picture Exchange Communication System (PECS)	3	3	3	3	3	3	3	3	3	3	3
High-P request sequence/Behavioral Momentum	2	2	2	2	2	2	2	2	2	2	2
Toilet Training	2	2	2	2	2	2	2	2	2	2	2
Time Out	1	1	1	1	1	1	1	1	1	1	1
Posture training vest with buzzer	1	1	1	1	1	1	1	1	1	1	1

Direct instruction	1	1	1	1	1	1	1	1	1	1	1
Average components per study	3.39	2.97	2.75	2.71	2.40	3.00	3.00	3.50	3.00	3.50	

Note: Includes all studies including those not meeting criteria for inclusion in outcomes analysis (N=128). COM = Communication, CB = Challenging Behavior, PR/A = Personal Responsibility/Adaptive Behavior, M = Motor, S/I = Social/InterpersonalL/SR = Learning/ School Readiness, A= Academic, P= Play, V = Vocational, Mix = Mixed

Table 3

WWC by skill domain

Skills Domain	Total	Meets without Reservation	Meets with Reservation	Does not Meet
Increasing	77	11	15	51
Academic	4	0	1	3
Communication	37	8	6	23
Learning/School Readiness	6	1	1	4
Communication and Learning/School Readiness	1	0	0	1
Motor	8	1	1	6
Personal Responsibility/Adaptive	12	1	3	8
Play	2	0	0	2
Social/Interpersonal	5	0	2	3
Vocational	2	0	1	1
Decreasing				
Challenging Behavior	32	7	6	19
Mixed	18	3	4	11
Communication and Challenging Behavior	8	2	2	4
Learning/School Readiness and Challenging Behavior	4	0	2	2
Motor and Challenging Behavior	1	0	0	1
Personal Responsibility/Adaptive and Challenging Behavior	4	1	0	3
Social/Interpersonal and Challenging Behavior	1	0	0	1
Total	127	21	25	81

Table 4

Tau-U Effect Sizes for Studies Included in Analysis

Study	Outcome Domain	Design	Design Quality	n with DS	Pairs	AB	Tau-U	90% CI	TAU-U Effect
Adamo, Wu, Wolery, Hemmeter, Ledford & Barton (2015)	Motor	ABAB (3)	MWR (2), MET (1)	3	248	6	0.91	[0.68,1.0]	Medium
Ayres & Cihak (2010)	Personal Responsibility /Adaptive	MBD	MWR	1	144	6	1	[0.89,1.0]	Large
Bauer & Jones (2015)	Communication	MPD (8)	MWR (6) DNM (2)	5	3445	18	0.93	[.74,1.0]	Large
Carbone, Sweeney-Kerwin, Attanasio & Kasper (2010)	Communication	MBD[a]	MET	1	170	1	1	[0.73,1.0]	Large
Cottrel, Montague, Farb & Throne (1980)	Communication	MBD	MET	1	1290	6	0.86	[0.70,1.0]	Medium
Davis, Brady, Williams & Hamilton (1992)	Learning/School Readiness	MBD	MWR	1	265	3	1	[0.73,1.0]	Large
Deleon et al. (2008)	Personal Responsibility /Adaptive	ABA(BCD)A (BCD)C(BC D)(BC)B(BE)(BCE)C	MWR	1	331	2	0.59	[0.25,0.92]	Small
Ducker & Moonen (1986)	Communication	MBD (	MWR	3	380	9	0.88	[0.69,1.0]	Medium
Feeley & Jones (2008)	Communication	MPD	MWR	1	298	3	0.76	[0.48,1.0]	Medium
Feeley, Jones, Blackburn, & Bauer (2011)	Communication	MBD (7)	MWR (2), DNM (5)	4	572	7	0.62	[0.44,0.8]	Small
Francisco & Hanley (2012)	Communication	MBD	MET	1	247	2	1	[0.72,1.0]	Large
Hagopian, Paclawskyj, & Kuhn (2005)	Challenging Behavior	ABAB	MET	1	102	2	0.98	[0.61,1.0]	Large
Haring & Kennedy (1990)	Challenging	ABCBCB (2)	MWR (2)	1	100	6	0.9	[.59,1.0]	Medium

Heller, Allgood, Ware, Arnold, & Casetelle (1996)	Behavior Communication	MPD[a]	MWR	1	15	1	1	[0.26,1.0]	Large
Hetzroni & Roth (2003)	Challenging Behavior	MPD[a]	MWR	1	16	1	1	[0.29,1.0]	Large
Hildebrand, Martin, Furer and Hazen (1990)	Vocational	MPD[a]	MET	1	342	1	0.86	[0.54,1.0]	Medium
Kettering, Neef, Kelley & Heward (2018)	Challenging Behavior, Communication	ABAB	MWR	1	114	4	.98	[0.67,1.0]	Large
Lafasakis & Sturmey (2007)	Communication, Learning/School Readiness	MPD[a]	MWR	1	71	2	0.33	[-0.1,0.77]	Small
Lalli, Browder, Mace, & Brown (1993)	Communication, Challenging Behavior	MBD[a]	MWR	1	64	1	1.0	[0.67,1.0]	Large
LeBlanc, Geiger, Sautter, & Sidener (2007)	Communication, Challenging Behavior	MPD[a]	MWR	1	120	2	0.65	[-0.27,1.0]	Medium
Mace, Hock, Lalli, West, Belfiore, Pinter and Brown (1988)	Learning/School Readiness, Challenging Behavior	ABABC	MET	2	90	2	1	[0.63,1.0]	Large
Martin, Rusch, James, Decker, Trtol (1982)	Personal Responsibility	MBD[a]	MET	1	264	1	1	[.61,1.0]	Large
Mazaleski, Iwata, Vollmer, Zarcone and Smit (1993)	Challenging Behavior	ABAB	MWR	1	490	2	1	[0.82,1.0]	Large
Mechling & Gast (1997)	Personal Responsibility /Adaptive	ABABAB	MWR (4)	4	184	18	1	[0.87,1.0]	Large
Morgan & Salzberg (1992)	Vocational	MPD[a]	MWR	1	498	4	0.98	[0.762,1.0]	Large
Poulson (1988)	Communication	ABAB (3)	MET (2) DNM (1)	3	951	5	0.81	[0.64,0.98]	Medium

Rapp, Vollmer, St. Peter, Dozier, & Cotnoir (2004)	Challenging Behavior	ABABA	MWR	1	402	8	0.25	[0.05, 0.44]	Small
Romaniuk, Miltenberger, Conyers, Jenner, Jurgens & Ringenberg (2002)	Challenging Behavior	ABABC	MWR	1	72	2	0.38	[-0.03,0.79]	Small
Rosine & Martine (1983)	Challenging Behavior	MBD[a]	MET	1	273	1	0.91	[0.50,1.0]	Medium
Schlosser, Belfiore, Nigam, Blischak & Hetzroni (1995)	Communication	MPD	MET	1	3193	8	0.67	[0.55,0.79]	Medium
Sigafoos, Reichle, Doss, Hall and Pettitt (1990)	Communication	MPD	MET (2)	2	719	4	0.88	[0.68,1.0]	Medium
Smith, Iwata, Vollmer, & Zarcone (1993)	Challenging Behavior	ABAB, ABCBC	MWR (2)	1	176	4	1	[0.63,1.0]	Large
Thompson et al. (2007)	Communication	ABAB	MWR	1	501	2	0.49	[0.15,0.82]	Small
Thompson, Iwata, Conners, & Roscoe (1999)	Challenging Behavior	ABAB	MET	1	390	2	0.77	[0.50,1.0]	Medium
Warren, Yoder, Gazdag, Kim & Jones (1993)	Communication	MBD	MET	1	1606	3	0.87	[0.691,1.0]	Medium
Werts, Hoffman and Darcy (2011)	Academic	MBD	MWR	1	354	3	0.77	[0.513,1.0]	Medium
Zarcone et al. (1993)	Learning/School Readiness, Challenging Behavior	MBD[a]	MWR	2	2114	4	0.65	[0.47,0.83]	Medium

Note. [a] Studies include a multiple baseline/multiple probe design across participants where not all of the participants in the design have DS. The effect sizes shown in this table represent a weighted average. N = number of participants, MBD = multiple-baseline design, MPD – multiple probe design, AB = number of phase contrasts, CI = confidence interval, DNM: did not meet WWC standards, MWR = met with WWC reservations, MET = met WWC without reservations.

Figures

Appendix A

PRISMA Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	3-6
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	6
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	6
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	8
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	7-8
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	7
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	8-9
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	9
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	9, 39-43

Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	10
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	10-11
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	11-12
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	12-13
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	n/a
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	33, Figure 1
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	13-14
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	15
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	14-15, 35, Figure 3
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	15-16, Figure 3
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	15, Figure 2
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	n/a
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	16-19
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	19-20

Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	20-21
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	N/A

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

Appendix B: Outcomes

Skill Domain	Definition
Increasing Behavior	
Social/Interpersonal	Skills needed to interact with others. The tasks comprising this category require social interaction with one or more individuals. Dependent variables associated with these tasks include but are not limited to joint attention, friendship, social and pretend play, social skills, social engagement, social problem solving, and appropriate participation in group activities. The area of pragmatics is not included in this list because it will be addressed in the communication section.
Academic	Performance on tasks typically taught and used in school settings. Examples include Reading, writing, mathematics.
Communication	Ability to express wants, needs, choices, feelings, or ideas. For example, Mand, tact, verbal imitation Communication tasks involve verbal or nonverbal signaling to a social partner regarding content of sharing of experiences, emotions, information, or affecting the partner's behavior, and behaviors that involve understanding a partner's intentional signals for the same purposes. This systematic means of communication involves the use of sounds or symbols. Dependent variables associated with these tasks include but are not restricted to requesting, labeling, receptive, conversation, greetings, nonverbal, expressive, syntax, speech, articulation, discourse, vocabulary, and pragmatics.
Learning/School Readiness	Performance during a task that is not directly related to task content (e.g. on task behavior). Learning readiness tasks serve as the foundation for successful mastery of complex skills in other domains identified. Dependent variables associated with these tasks include but are not restricted to imitation, following instructions, sitting skills, and attending to environmental sounds. Motor imitation is included in this category not motor skills.
Motor	Movement or motion, including both fine and gross motor skills, or related to sensory system/sensory functioning. Motor skills involve tasks that require coordination of muscle systems to produce a specific goal involving either fine motor or gross motor skills or visual-motor coordination. Fine motor skills require manipulation of objects using precise movements to produce the desired outcome. Examples of fine motor skills include but are not restricted to cutting, coloring, writing, typing, and threading beads. Gross motor skills involve larger muscle movements and include but are not restricted

	to sitting, standing, walking, and throwing/catching balls.
Personal Responsibility/Adaptive	Independent living skills and personal care skills (e.g. washing, toileting) This category targets tasks that involve activities embedded in everyday routines. Dependent variables associated with these tasks include but are not restricted to feeding, sleeping, dressing, toileting, cleaning, family and/or community activities, health and fitness, phone skills, time and money management, and self-advocacy.
Play	Use of toys or leisure materials Play tasks involve non-academic and non-work-related activities that do not involve self-stimulatory behavior or require interaction with other persons. Dependent variables associated with these tasks may include but are not restricted to functional independent play (i.e., manipulation of toys to determine how they "work" or appropriate use of toys that do not involve pretense, games). Whenever social play was targeted (independently or in conjunction with make-believe play), it was placed in the "interpersonal" categories.
Vocational	Employment or employment preparation or relate to technical skills required for a specific job
Decreasing Behavior	
Challenging Behavior	Decreasing or eliminating behaviors that interfere with the individual's ability to learn.

Note: Table adapted from ONTABA Scientific Expert Task Force for Treatment of Autism Spectrum Disorders (OSETT-ASD) (2017). Evidence-based practices for individuals with autism spectrum disorder: Recommendations for caregivers, practitioners, and policy makers. Toronto, ON. Retrieved from http://www.ontaba.org/pdf/ONTABA%20OSETT-ASD%20REPORT%20WEB.pdf

Appendix C: Articles Included in the Analysis

Adamo, E.K., Wu, J., Wolery, M., Hemmeter, M.L., Ledford, J.R., & Barton, E.E. (2015). Using video modeling, prompting, and behavior-specific praise to increase moderate-to-vigorous physical activity for young children with Down syndrome. *Journal of Early Intervention, 37*(34), 270-285.

Anderson, C.M., Doughty, S.S., Doughty, A.H., Williams, D.C., & Saunders, K.J. (2010). Evaluation of stimulus control over a communication response as an intervention for stereotypical responding. *Journal of Applied Behavior Analysis, 43*(2), 333-339.

Apolloni, T., Cooke, S.A., & Cooke, T.P. (1977). Establishing a normal peer as a behavioral model for developmentally delayed toddlers. *Perceptual and Motor Skills, 44,* 231-241.

Ayres, K., & Cihak, D. (2010). Computer-and video-based instruction of food-preparation skills: Acquisition, generalization, and maintenance. *Intellectual and Developmental Disabilities, 48*(3), 195-208.

Barmann, B.C. (2013). The use of overcorrection with artificial nails in the treatment of chronic fingernail biting. *Mental Retardation, 17,* 309-317.

Bauer, S. M., & Jones, E. A. (2015). Requesting and verbal imitation intervention for infants with Down syndrome: Generalization, intelligibility, and problem solving. *Journal of Developmental and Physical Disabilities, 27*(1), 37-66.

Bauer, S.M., Jones, E.A., & Feeley, K.M. (2014). Teaching responses to questions to young children with Down syndrome. *Behavioral Interventions, 29*(1), 36-49.

Bennett, F., Eisenman, P., French R., Henderson, H., & Shultz, B. (1989). The effect of a token economy on the exercise behavior of individuals with Down syndrome. *Adapted Physical Activity Quarterly, 6*(3), 230-246.

Belisle, J., Stanley, C. R., Alholail, A. M., Galliford, M. E. & Dixon, M. R. (2019). Abstraction of tactile properties by individuals with autism and Down syndrome using a picture‐based communication system. *Journal of Applied Behavior Analysis, 52*(2), 467-475.

Benoit, D. A., Edwards, R. P., Olmi, D. J., Wilczynski, S. M., & Mandal, R. L. (2001). Generalization of a positive treatment package for child noncompliance. *Child & Family Behavior Therapy, 23*(2), 19-32.

Binger, C., & Light, J. (2007). The effect of aided AAC modeling on the expression of multi-symbol messages by preschoolers who use AAC. *Augmentative and Alternative Communication, 23*(1), 30-43.

Bosner, S. M., & Belfiore, P. J. (2001). Strategies and considerations for teaching an adolescent with Down syndrome and type I diabetes to self-administer insulin. *Education and Training in Mental Retardation and Developmental Disabilities*, 94-102.

Brooks, A., Todd, A. W., Tofflemoyer, S., & Horner, R. H. (2003). Use of functional assessment and a self-management system to increase academic engagement and work completion. *Journal of Positive Behavior Interventions, 5*(3), 144-152.

Boutot, E. & DiGangi, S. (2018). Effects of activation of preferred stimulus on tummy time behavior of an infant with Down syndrome and associated hypotonia. *Behavior Analysis in Practice, 11*(2), 144-147.

Bruder, M.B. (1987). Parent-to-parent teaching. *American Journal of Mental Deficiency, 91,* 435-438.

Carbone, V.J., & Sweeney-Kerwin, E.J. (2010). Increasing the vocal responses of children with autism and developmental disabilities using manual sign mand training and prompt delay. *Journal of Applied Behavior Analysis, 43,* 705-709.

Carr, J. (1987). Bedwetting: A new approach to treatment in a mental handicapped boy. *Child: Care, Health and Development, 13,* 239-245.

Chambers, M., & Rehfeldt, R. A. (2003). Assessing the acquisition and generalization of two mand forms with adults with severe developmental disabilities. *Research in Developmental Disabilities, 24*(4), 265-280.

Coe, D., Matson, J., Fee, V., Manikam, R., & Linarello, C. (1990). Training nonverbal and verbal play skills to mentally retarded and autistic children. *Journal of Autism and Developmental Disorders, 20*(2), 177-187.

Cole, C. L., & Levinson, T. R. (2002). Effects of within-activity choices on the challenging behavior of children with severe developmental disabilities. *Journal of Positive Behavior Interventions, 4*(1), 29-37.

Cottrell, A. W., Montague, J., Farb, J., & Throne, J. M. (1980). An operant procedure for improving vocabulary definition performances in developmentally delayed children. *Journal of Speech and Hearing Disorders, 45*(1), 90-102.

Davis, C.A., Brady, M.P., Williams, R.E., & Hamilton, R. (1992). Effects of high-probability on the acquisition and generalization of responses to requests in young children with behavior disorders. *Journal of Applied Behavior Analysis, 25*(4), 905-916.

DeLeon, I. G., Hagopian, L. P., Rodriguez Catter, V., Bowman, L. G., Long, E. S., & Boelter, E. W. (2008). Increasing wearing of prescription glasses in individuals with mental retardation. *Journal of Applied Behavior Analysis, 41*(1), 137-142.

Desrosiers, G. (1987). Emergency telephone calls: A multiple baseline study using instructional programming. *Journal of Behavior Therapy and Experimental Psychiatry, 18*(4), 373-380.

Doughty, S. S., Anderson, C. M., Doughty, A. H., Williams, D. C., & Saunders, K. J. (2007). Discriminative control of punished stereotyped behavior in humans. *Journal of the Experimental Analysis of Behavior, 87*(3), 325-336.

Ducker, P.C., & Moonen, X.M. (1986). The effects of two procedures on spontaneous signing with Down's Syndrome children. *Journal of Intellectual Disability Research*, *30*(4), 355-364.

Duker, P. C., & Michielsen, H. M. (1983). Cross-setting generalization of manual signs to verbal instructions with severely retarded children. *Applied Research in Mental Retardation, 4*(1), 29-40.

Dunn, J., Lockwood, K., Williams, D.E., & Peacock, S. (1997). A seven-year follow-up of treating rumination with dietary satiation. *Behavioral Interventions*, *12*(4), 163-172.

Erbas, D., Ozen, A., & Acar, C. (2004). Comparison of two approaches for identifying reinforces in teaching figure coloring to students with Down syndrome. *Education and Training in Developmental Disabilities, 39,* 253- 264.

Farb, J., & Throne, J.M. (1978). Improving the generalized mnemonic performance of a Down's Syndrome child. *Journal of Applied Behavior Analysis, 11*(3), 413-419.

Feeley, K., & Jones, E. (2007). Teaching spontaneous responses to a young child with Down syndrome. *Down Syndrome Research and Practice, 12*(2), 148-152.

Feeley, K., & Jones, E. (2008). Strategies to address challenging behavior in young children with Down syndrome. *Down Syndrome Research and Practice, 12,* 153-163.

Feeley, K., Jones, E., Blackburn, C., Bauer, S. (2011). Advancing imitation and requesting skills in toddlers with Down syndrome. *Research in Developmental Disabilities, 32*(6), 2415-2430.

Fisher, W. W., Bowman, L. G., Thompson, R. H., Contrucci, S. A., Burd, L., & Alon, G. (1998). Reductions in self injury produced by transcutaneous electrical nerve stimulation. *Journal of Applied Behavior Analysis, 31*(3), 493-496.

Fisher, W.W., Piazza, C.C., Bowman, L.G., Hanley, G.P., & Adelinis, J.D. (1997). Direct and collateral effects of restraints and restraint fading. *Journal of Applied Behavior Analysis, 30*(1), 105-120.

Fox, J., Shores, R., Lindeman, D., & Strain, P. (1986). Maintaining social initiations of withdrawn handicapped and nonhandicapped preschoolers through a response-dependent fading tactic. *Journal of Abnormal Child Psychology, 14*(3), 387-396.

Francisco, M.T., & Hanley, G.P. (2012). An evaluation of progressively increasing intertrial intervals on the acquisition and generalization of three social skills. *Journal of Applied Behavior Analysis, 45*(1), 137-142.

Gardner, W.I., Cole, C.L., Berry, D.L., & Nowinski, J.M. (1983). Reduction of disruptive behaviors in mental retarded adults: A self-management approach. *Behavior Modification, 7*(1), 76-96.

Gibbons, B.G., Williams, K.E., & Riegel, K.E. (2007). Reducing tube feeds and tongue trust: Combining an oral-motor and behavioral approach to feeding. *The American Journal of Occupational Therapy, 61,* 394-401.

Goldstein, H., English, K., Shafer, K., & Kaczmarek, L. (1997). Interaction among preschoolers with and without disabilities: Effects of across-the-day peer intervention. Journal of *Speech, Language, and Hearing Research, 40*(1), 33-48.

Griffen, A.K., Wolery, M., & Schuster, J.W. (1992). Triadic instruction of chained food preparation responses: Acquisition and observational learning. *Journal of Applied Behavior Analysis, 25*(1), 193-204.

Hagopian, L. P., Paclawskyj, T. R., & Kuhn, S. C. (2005). The use of conditional probability analysis to identify a response chain leading to the occurrence of eye poking. *Research in Developmental Disabilities, 26*(4), 393-397.

Hall, A. M., Neuharth-Pritchett, S., & Belfiore, P. J. (1997). Reduction of aggressive behaviors with changes in activity: Linking descriptive and experimental analyses. *Education and Training in Mental Retardation and Developmental Disabilities*, 331-339.

Hanson, M. J., & Hanline, M. F. (1985). An analysis of response-contingent learning experiences for young children. *Journal of the Association for Persons with Severe Handicaps, 10*(1), 31-40.

Haring, T. G. (1985). Teaching between-class generalization of toy play behavior to handicapped children. *Journal of Applied Behavior Analysis, 18*(2), 127-139.

Haring, T. G., Breen, C. G., & Laitinen, R. E. (1989). Stimulus class formation and concept learning: Establishment of within-and between-set generalization and transitive relationships via conditional discrimination procedures. *Journal of the Experimental Analysis of Behavior, 52*(1), 13-25.

Haring, T.G., & Kennedy, C.H. (1990). Contextual control of problem behavior in students with severe disabilities. *Journal of Applied Behavior Analysis, 23,* 235-243.

Haring, T.G., Roger, B., Lee, M., Breen, C., & Ross, R.G. (1986). Teaching social language to moderately handicapped students. *Journal of Applied Behavior Analysis, 19*(2), 159-171.

Heller, K. W., Allgood, M. H., Ware, S., Arnold, S. E., & Castelle, M. D. (1996). Initiating requests during community-based vocational training by students with mental retardation and sensory impairments. *Research in Developmental Disabilities, 17*(3), 173-184.

Hemmeter, M. L., Ault, M. J., Collins, B. C., & Meyer, S. (1996). The effects of teacher-implemented language instruction within free time activities. *Education & Training in Mental Retardation & Developmental Disabilities, 31*(3), 203-212

Hetzroni, O. E., & Roth, T. (2003). Effects of a positive support approach to enhance communicative behaviors of children with mental retardation who have challenging behaviors. *Education and Training in Developmental Disabilities,* 95-105.

Hewitt, L. E., Hinkle, A. S., & Miccio, A. W. (2005). Intervention to improve expressive grammar for adults with Down syndrome. *Communication Disorders Quarterly, 26*(3), 144-155.

Hildebrand, R.G., Martin G.L., Furer, P., & Hazen, A. (1990). A recruitment-of-praise package to increase productivity levels of developmentally handicapped workers. *Behavior Modification, 14*(1), 97-113.

Huguenin, N.H. (1993). Reducing chronic noncompliance in an individual with severe mental retardation to facilitate community integration. *Mental Retardation, 51*(5), 332-339.

Huynen, K.B., Lutzker, J.R., Bigelow, K.M., Touchette, P.E., & Campbell, R.V. (1996). Planned activities training for mothers of children with developmental disabilities: Community generalization and follow-up. *Behavior Modification, 20*(4), 406-427.

Iacono, T., & Duncum, J. (1995). Comparison of sign alone and in combination with an electronic communication device in early language intervention: Case study. *Augmentative and Alternative Communication, 11*(4), 249-259.

Irwin, K.C. (1991). Teaching children with Down syndrome to add by counting-on. *Education and Treatment of Children, 14*(2), 128-142.

Jones, E., Feeley, K., & Blackburn, C. (2010). A preliminary study of intervention addressing early developing requesting behaviors in young infants with Down syndrome. *Down Syndrome Research and Practice, 12*(2), 98-102.

Kennedy, C. H., & Haring, T. G. (1993). Combining reward and escape DRO to reduce the problem behavior of students with severe disabilities. *Journal of the Association for Persons with Severe Handicaps, 18*(2), 85-92.

Kettering, T. L., Neef, N. A., Kelley, M. E. &Heward, W. L. (2018). A comparison of procedures for unpairing conditioned reflexive motivating operations. *Journal of the Experimental Analysis of Behavior, 109*(2), 422-432.

Konczak, L.J., & Johnson, C.M. (1983). Reducing inappropriate verbalizations in a sheltered workshop through differential reinforcement of other behavior, *Education and Training of the Mentally Retarded,* 120-124.

Kouri, T. (1988). Effects of simultaneous communication in a child-directed treatment approach with preschoolers with severe disabilities. *Augmentative and Alternative Communication, 4*(4), 222-232.

Lafasakis, M., & Sturmey, P. (2007). Training parent implementation of discrete-trial teaching: Effects on generalization of parent teaching and child correct responding. *Journal of applied behavior analysis, 40*(4), 685-689.

Lalli, J. S., Browder, D. M., Mace, F. C., & Brown, D. K. (1993). Teacher use of descriptive analysis data to implement interventions to decrease students' problem behaviors. *Journal of Applied Behavior Analysis, 26*(2), 227-238.

Lalli, J.S., Mace, F.C., Browder, D., & Brown, D.K. (1989). Comparison of treatments to teach number matching skills to adults with moderate mental retardation. *Mental Retardation, 27*(2), 75-85.

Lancioni, G. E., Singh, N. N., O'Reilly, M. F., Sigafoos, J., Oliva, D., Smaldone, A., & La Martire, M. L. (2009). Two persons with multiple disabilities use a mouth-drying response to reduce the effects of their drooling. *Research in Developmental Disabilities, 30*(6), 1229-1236.

Lancioni, G. E., Van Bergen, I., & Furniss, F. (2002). Urine alarms and prompts for fostering daytime urinary continence in a student with multiple disabilities: A replication study. *Perceptual and motor skills, 94*(3), 867-870.

LeBlanc, L. A., Geiger, K. B., Sautter, R. A., & Sidener, T. M. (2007). Using the Natural Language Paradigm (NLP) to increase vocalizations of older adults with cognitive impairments. *Research in Developmental Disabilities, 28*(4), 437-444.

Lohrmann-O'Rourke, S., & Yurman, B. (2001). Naturalistic assessment of and intervention for mouthing behaviors influenced by establishing operations. *Journal of Positive Behavior Interventions, 3*(1), 19-27.

Luciano, C., Barnes-Holmes, Y., & Barnes-Holmes, D. (2002). Establishing reports of saying and doing and discriminations of say–do relations. *Research in Developmental Disabilities, 23*(6), 406-421.

Luciano-Soriano, M. C., Molina-Cobos, F. J., & Gómez-Becerra, I. (2000). Say-do-report training to change chronic behaviors in mentally retarded subjects. *Research in Developmental Disabilities*, *21*(5), 355-366.

Luiselli, J.K. (1991). Functional assessment and treatment of self-injury in a pediatric, nursing-case resident. *Behavioral Interventions, 6,* 311-319.

Mace, F.C., Hock, M.L., Lalli, J.S., West, B.J., Belfiore, P., & Pinter, E. (1988). Behavioral momentum in the treatment of noncompliance. *Journal of Applied Behavior Analysis, 21*(2), 123-141.

Magee, S. K., & Ellis, J. (2001). The detrimental effects of physical restraint as a consequence for inappropriate classroom behavior. *Journal of Applied Behavior Analysis, 34*(4), 501-504.

Marcus, B. A., & Vollmer, T. R. (1995). Effects of differential negative reinforcement on disruption and compliance. *Journal of Applied Behavior Analysis, 28*(2), 229-230.

Marcus, B.A., & Vollmer, T.R. (1996). Combining noncontingent reinforcement and differential reinforcement schedules as treatment for aberrant behavior. *Journal of Applied Behavior Analysis, 29*(1), 43-51.

Martella, R.C., Marchand-Martella, N.E., Young, K.R., & Macfarlane, C.A. (1995). Determining the collateral effects of peer tutor training on a student with severe disabilities. *Behavior Modification, 19*(2), 170-191.

Martin, J.E., Rusch, F.R., James, V.L., Decker, P.J., & Trtol, K.A. (1982). The use of picture cues to establish self-control in the preparation of complex meals by mentally retarded adults. *Applied Research in Mental Retardation, 3*(2), 105-119.

Matthews, J.R., Hodson, G.D., Crist, W.G., & LaRoche, G.R. (1992). Teaching young children to use contact lenses. *Journal of Applied Behavior Analysis, 25*(1), 229-235.

Mazaleski, J.L., Iwata, B.A., Vollmer, T.R., Zarcone, J.R., & Smith, R.G. (1993). Analysis of the reinforcement and extinction components in DRO contingencies with self-injury. *Journal of Applied Behavior Analysis, 26*(2), 143-156.

McBride, B. J., & Schwartz, I. S. (2003). Effects of teaching early interventionists to use discrete trials during ongoing classroom activities. *Topics in Early Childhood Special Education, 23*(1), 5-17.

McComas, J. J., Thompson, A., & Johnson, L. (2003). The effects of presession attention on problem behavior maintained by different reinforcers. *Journal of Applied Behavior Analysis, 36*(3), 297-307.

McDonald, M.R., & Budd, K.S. (1983). "Booster Shots" following didactic parent training: Effects of follow-up using graphic feedback and instructions. *Behavior Modification, 7*(2), 211-223.

McDonnell, A. P. (1996). The acquisition, transfer, and generalization of requests by young children with severe disabilities. *Education and Training in Mental Retardation and Developmental Disabilities*, 213-234.

Mechling, L. C., & Gast, D. L. (1997). Combination audio/visual self-prompting system for teaching chained tasks to students with intellectual disabilities. *Education and Training in Mental Retardation and Developmental Disabilities, 32*, 138-153.

Mirenda, P., & Dattilo, J. (1987). Instructional techniques in alternative communication for students with severe intellectual handicaps. *Augmentative and Alternative Communication, 3*(3), 143-152.

Morgan, R. L., & Salzberg, C. L. (1992). Effects of video-assisted training on employment-related social skills of adults with severe mental retardation. *Journal of Applied Behavior Analysis, 25*(2), 365-383.

Neil, N., & Jones, E.A. (2015). Studying treatment intensity: Lessons from two preliminary studies. *Journal of Behavioral Education, 24*(1), 51-73.

Neil, N., & Jones, E.A. (2016). Repetitive behavior in children with Down syndrome: Functional analysis and intervention. *Journal of Developmental and Physical Disabilities, 28*(2), 267-288.

Olenick, D.L., & Pear, J.J. (1980). Differential reinforcement of correct responses to probes and prompts in picture-name training with severely retarded children. *Journal of Applied Behavior Analysis, 13*(1), 77-89.

Oliver, C., Hall, S., Hales, J., Murphy, G., & Watts, D. (1998). The treatment of severe self-injurious behavior by the systematic fading of restraints: Effects on self-injury, self-restraint, adaptive behavior and behavioral correlates of affect. *Research in Developmental Disabilities, 19*(2), 143-165.

O'Reilly, M. F., Murray, N., Lancioni, G. E., Sigafoos, J., & Lacey, C. (2003). Functional analysis and intervention to reduce self-injurious and agitated behavior when removing protective equipment for brief time periods. *Behavior modification, 27*(4), 538-559.

O'Toole, C., Lee, A. S. Y., Gibbon, F. E., van Bysterveldt, A. K., & Hart, N. J. (2018). Parent‐mediated interventions for promoting communication and language development in young children with Down syndrome. Cochrane Database of Systematic Reviews, (10). DOI: 10.1002/14651858.CD012089.pub2

Polvinale, R.A., & Lutzker, J.R. (1980). Elimination of assaultive and inappropriate sexual behavior by reinforcement and social-restitution. *Mental Retardation, 18*(1), 27-31.

Poulson, C.L. (1988). Operant conditioning of vocalization rate of infants with Down syndrome. *American Journal of Mental Retardation, 93,* 57-63.

Pufpaff, L. A., Blischak, D. M., & Lloyd, L. L. (2000). Effects of modified orthography on the identification of printed words. *American Journal on Mental Retardation, 105*(1), 14-24.

Purdy, A.H., Deitz, J.C., & Harris, S.R. (1987). Efficacy of two treatment approaches to reduce tongue protrusion of children with Down syndrome. *Developmental Medicine and Child Neurology, 29*(4), 469-476.

Rapp, J. T., Vollmer, T. R., St Peter, C., Dozier, C. L., & Cotnoir, N. M. (2004). Analysis of response allocation in individuals with multiple forms of stereotyped behavior. *Journal of Applied Behavior Analysis, 37*(4), 481-501.

Remington, B., & Clarke, S. (1993). Simultaneous communication and speech comprehension. Part II: Comparison of two methods of overcoming selective attention during expressive sign training. *Augmentative and Alternative Communication, 9*(1), 49-60.

Remington, B., & Clarke, S. (1993). Simultaneous communication and speech comprehension. Part I: Comparison of two methods of teaching expressive signing and speech comprehension skills. *Augmentative and Alternative Communication, 9*(1), 36-48.

Repp, A.C., & Karsh, K.G. (1994). Hypothesis-based interventions for tantrum behaviors of persons with developmental disabilities in school settings. *Journal of Applied Behavior Analysis, 27*(1), 21-31.

Richling, S.M., Rapp, J.T., Carroll, R.A., Smith, J.N., Nystedt, A., & Siewert, B. (2011). Using noncontingent reinforcement to increase compliance with wearing prescription prostheses. *Journal of Applied Behavior Analysis, 44*(2), 375-379.

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*(4), 349-362.

Rosales, R., & Rehfeldt, R. A. (2007). Contriving transitive conditioned establishing operations to establish derived manding skills in adults with severe developmental disabilities. *Journal of Applied Behavior Analysis, 40*(1), 105-121.

Rosine, L. P., & Martin, G. L. (1983). Self-management training to decrease undesirable behavior of mentally handicapped adults. *Rehabilitation Psychology, 28*(4), 195.

Rudrud, E.H., Ziarnik, J.P., & Colman, G. (1984). Reduction of tongue protrusion of a 24-year-old woman with Down syndrome through self-monitoring. *American Journal of Mental Deficiency, 88,* 647-652.

Salzberg, C.L., & Villani, T.V. (1983). Speech training by parents of Down syndrome toddlers: Generalization across settings and instructional contexts. *American Journal of Mental Deficiency, 87,* 403-413.

Schleien, S. J., Wehman, P., & Kiernan, J. (1981). Teaching leisure skills to severely handicapped adults: An age-appropriate darts game. *Journal of Applied Behavior Analysis, 14(*4), 513-519.

Schlosser, R. W., Belfiore, P. J., Nigam, R., Blischak, D., & Hetzroni, O. (1995). The effects of speech output technology in the learning of graphic symbols. *Journal of Applied Behavior Analysis, 28*(4), 537-549.

Schmid, T. L. (1988). A comparison of two behaviour reduction procedures: Traditional extinction alone and interpolated reinforcement followed by extinction. *Journal of Intellectual Disability Research, 32*(4), 289-300.

Shaw, R., & Simms, T. (2009). Reducing attention-maintained behavior through the use of positive punishment, differential reinforcement of low rates, and response marking. *Behavioral Interventions, 24*(4), 249-263.

Shore, B. A., LeBlanc, D., & Simmons, J. (1999). Reduction of unsafe eating in a patient with esophageal stricture. *Journal of Applied Behavior Analysis, 32*(2), 225-228.

Sigafoos, J., Doss, S., & Reichle, J. (1989). Developing mand and tact repertoires in person with severe developmental disabilities using graphic symbols. *Research in Developmental Disabilities, 10*(2), 183-200.

Sigafoos, J., Reichle, J., Doss, S., Hall, K., & Pettitt, L. (1990). "Spontaneous" transfer of stimulus control from tact to mand contingencies. *Research in Developmental Disabilities, 11*(2), 165-176.

Smith, J., Hand, L., & Dowrick, P.W. (2014). Video feedforward for rapid learning of a picture-based communication system. *Journal of Autism and Developmental Disorders, 44*(4), 926-936.

Smith, R. G., Iwata, B. A., Vollmer, T. R., & Zarcone, J. R. (1993). Experimental analysis and treatment of multiply controlled self-injury. *Journal of Applied Behavior Analysis, 26*(2), 183-196.

Solnick, J.V., Rincover, A., & Peterson, C.R. (1977). Some determinants of the reinforcing and punishing effects of timeout. *Journal of Applied Behavior Analysis, 10*(3), 415-424.

Tarpley, H.D., & Schroeder, S.R. (1979). Comparison of DRO and DRI on rate of suppression of self-injurious behavior. *American Journal of Mental Deficiency, 84,* 188-194.

Tekin, E., & Kircaali-Iftar, G. (2002). Comparison of the effectiveness and efficiency of two response prompting procedures delivered by sibling tutors. *Education and Training in Mental Retardation and Developmental Disabilities*, 283-299.

Tekin-Iftar, E. (2003). Effectiveness of peer delivered simultaneous prompting on teaching community signs to students with developmental disabilities. *Education and Training in Developmental Disabilities,* 77-94.

Thompson, R. H., Iwata, B. A., Conners, J., & Roscoe, E. M. (1999). Effects of reinforcement for alternative behavior during punishment of self-injury. *Journal of Applied Behavior Analysis, 32*(3), 317-328.

Thompson, R.H., Cotnoir-Bichelman, N.M., McKerchar, P.M., Tate, T.L., & Dancho, K.A. (2007). Enhancing early communication through infant sign training. *Journal of Applied Behavior Analysis, 40*(1), 15-23.

Tiller, J., Stygar, M.K., Hess, C., & Reimer, L. (1982). Treatment of functional chronic stooped posture using a training device and behavior therapy: A case report. *Physical Therapy, 62,* 1597-1600.

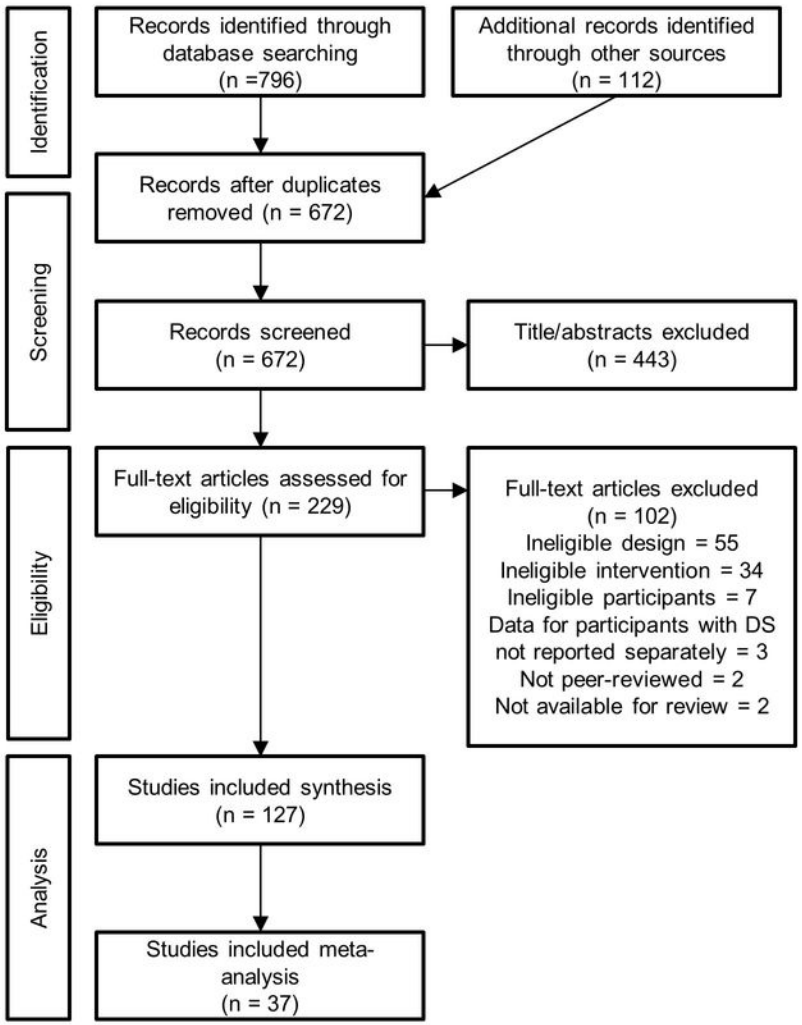
Wallace, A.M., Bechtel, D.R., Heatter, S., & Barry, L.M. (2016). A comparison of prompting strategies to teach intraverbals to an adolescent with Down syndrome. *The Analysis of Verbal Behavior, 32*(2), 225-232.

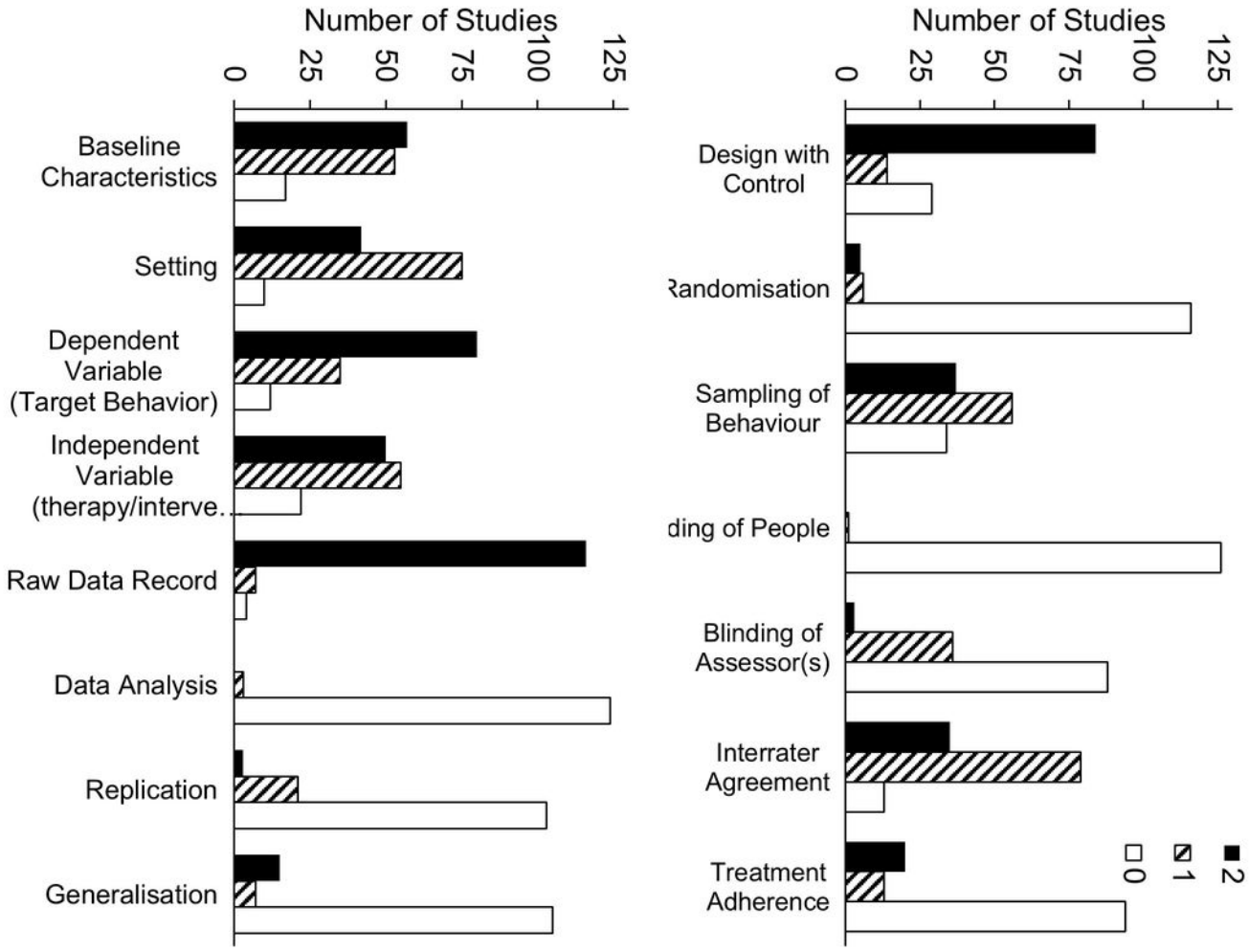
Warren, S. F., Yoder, P. J., Gazdag, G. E., Kim, K., & Jones, H. A. (1993). Facilitating prelinguistic communication skills in young children with developmental delay. *Journal of Speech, Language, and Hearing Research, 36*(1), 83-97.

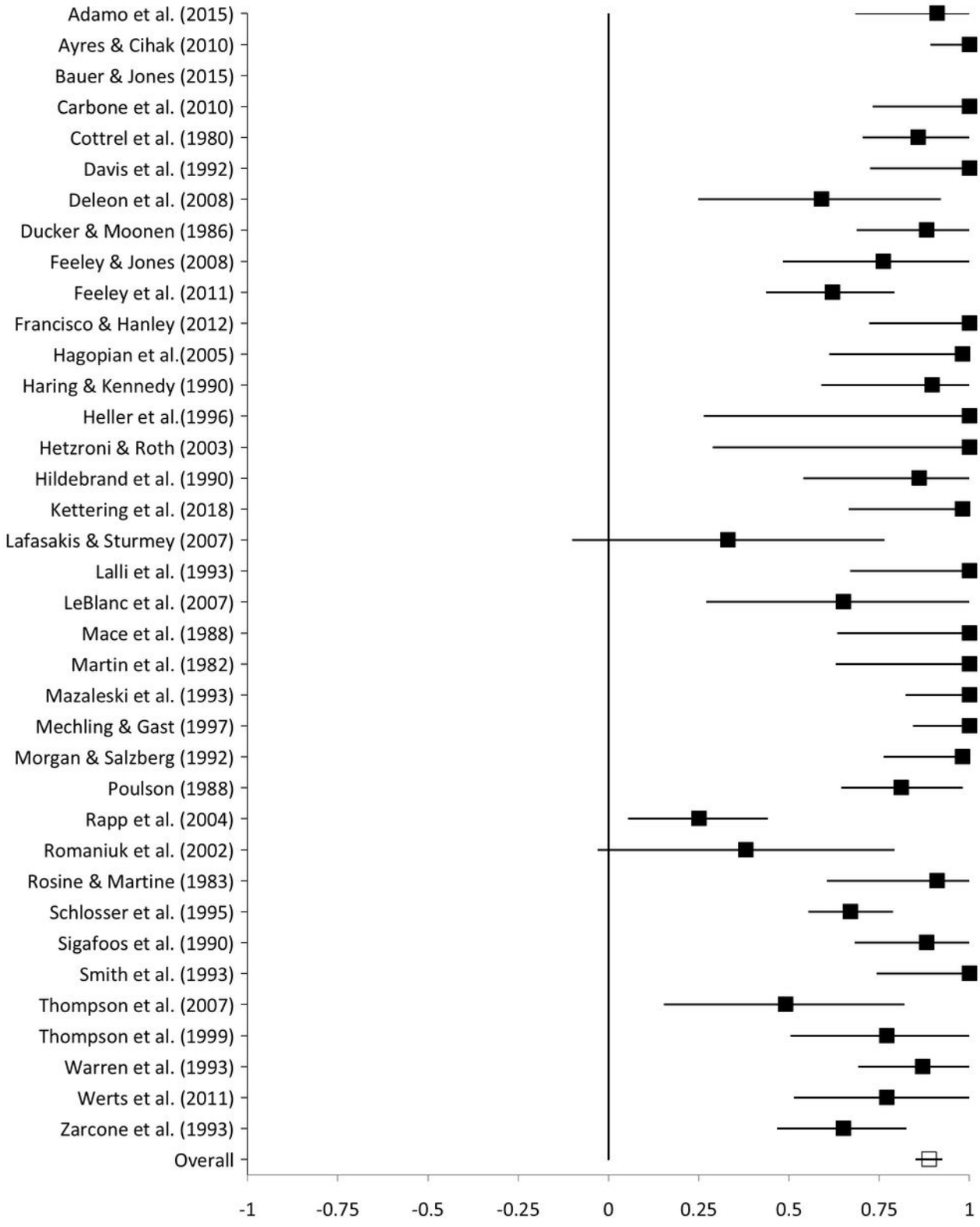
Werts, M.G., Hoffman, E.M., & Darcy, C. (2011). Acquisition of instructive feedback: Relation to target stimulus. *Education and Training in Autism and Developmental Disabilities, 46,* 134-149.

Winborn-Kemmerer, L., Wacker, D.P., Harding, J., Boelter, E., Berg, W., & Lee, J. (2010). Analysis of mand selection across different stimulus conditions. *Education and Treatment of Children, 33*(1), 49-64.

Zarcone, J.R., Iwata, B.A., Vollmer, T.R., Jagtiani, S., Smith, R.G., & Mazaleski, J.L. (1993). Extinction of self-injurious escape behavior with and without instructional fading. *Journal of Applied Behavior Analysis, 26*(3), 353-360.


Identification
Screening
Eligibility
Analysis
Records identified through database searching (n =796)
Additional records identified through other sources (n = 112)
Records after duplicates removed (n = 672)
Records screened (n = 672)
Title/abstracts excluded (n = 443)
Full-text articles assessed for eligibility (n = 229)
Full-text articles excluded (n = 102)
Ineligible design = 55
Ineligible intervention = 34
Ineligible participants = 7
Data for participants with DS not reported separately = 3
Not peer-reviewed = 2
Not available for review = 2
Studies included synthesis (n = 127)
Studies included meta-analysis (n = 37)


Number of Studies
0
25
50
75
100
125
Baseline Characteristics
Setting
Dependent Variable (Target Behavior)
Independent Variable (therapy/interve.
Raw Data Record
Data Analysis
Replication
Generalisation
Number of Studies
0
25
50
75
100
125
Design with Control
Randomisation
Sampling of Behaviour
ding of People
Blinding of Assessor(s)
Interrater Agreement
Treatment Adherence
2
1
0


Adamo et al. (2015)
Ayres & Cihak (2010)
Bauer & Jones (2015)
Carbone et al. (2010)
Cottrel et al. (1980)
Davis et al. (1992)
Deleon et al. (2008)
Ducker & Moonen (1986)
Feeley & Jones (2008)
Feeley et al. (2011)
Francisco & Hanley (2012)
Hagopian et al.(2005)
Haring & Kennedy (1990)
Heller et al.(1996)
Hetzroni & Roth (2003)
Hildebrand et al. (1990)
Kettering et al. (2018)
Lafasakis & Sturmey (2007)
Lalli et al. (1993)
LeBlanc et al. (2007)
Mace et al. (1988)
Martin et al. (1982)
Mazaleski et al. (1993)
Mechling & Gast (1997)
Morgan & Salzberg (1992)
Poulson (1988)
Rapp et al. (2004)
Romaniuk et al. (2002)
Rosine & Martine (1983)
Schlosser et al. (1995)
Sigafoos et al. (1990)
Smith et al. (1993)
Thompson et al. (2007)
Thompson et al. (1999)
Warren et al. (1993)
Werts et al. (2011)
Zarcone et al. (1993)
Overall
-1
-0.75
-0.5
-0.25
0
0.25
0.5
0.75
1