
Janine Montgomery  
*St. Amant Research Centre and University of Manitoba, Janine.Montgomery@umanitoba.ca*

Toby Martin  
*St. Amant Research Centre and University of Manitoba*

Shahin Shooshtari  
*St. Amant Research Centre and University of Manitoba*

Brenda M. Stoesz  
*St. Amant Research Centre and University of Manitoba*

Dustin J. Heinrichs  
*St. Amant Research Centre and University of Manitoba*

*See next page for additional authors*

**Abstract**  
This synthesis paper summarizes research literature addressing challenging behaviours in children and youth with autism spectrum disorders and developmental disabilities in school settings. We conducted a
comprehensive literature review to identify relevant peer-reviewed articles published between the years 2000 and 2011. The methodological quality of all studies was screened following a standard quality assessment checklist. Intervention effects were assessed quantitatively. Results of quality analysis and intervention effectiveness were integrated to identify interventions with strong support and evidence of effectiveness. We discuss results in terms of implications for intervention choice and implementation in school settings, limitations, and directions for future research.

ISSN 1918-5227
Pages 2-21

Follow this and additional works at: https://ir.lib.uwo.ca/eei
Part of the Education Commons

Cover Page Footnote
This project was funded by the Canadian Institute of Health Research. We wish to thank Drs. J. Virues Ortega and C. T. Yu and two anonymous reviewers for their reviews and suggestions for this paper.

Authors
Janine Montgomery, Toby Martin, Shahin Shooshtari, Brenda M. Stoesz, Dustin J. Heinrichs, Sebastian North, Lindsay Dodson, Quinn Senkow, and Joyce Douglas

This article is available in Exceptionality Education International: https://ir.lib.uwo.ca/eei/vol23/iss1/2

Janine Montgomery, Toby Martin, Shahin Shooshtari, Brenda M. Stoesz, Dustin J. Heinrichs
St. Amant Research Centre and University of Manitoba
Sebastian North
Pembina Trails School Division
Lindsay Dodson
St. Amant Research Centre and University of Manitoba
Quinn Senkow
Seine River School Division
Joyce Douglas
St. Amant School
& the KATYDID team

Abstract

This synthesis paper summarizes research literature addressing challenging behaviours in children and youth with autism spectrum disorders and developmental disabilities in school settings. We conducted a comprehensive literature review to identify relevant peer-reviewed articles published between the years 2000 and 2011. The methodological quality of all studies was screened following a standard quality assessment checklist. Intervention effects were assessed quantitatively. Results of quality analysis and intervention effectiveness were integrated to identify interventions with strong support and evidence of effectiveness. We discuss results in terms of implications for intervention choice and implementation in school settings, limitations, and directions for future research.

Challenging behaviours are intense, frequent, and/or long lasting and disrupt learning, threaten physical safety, and/or lead to restricted access to community facilities and services (e.g., education; Emerson et al., 2001). These debilitating behaviours (see examples below) are commonly demonstrated by individuals with communication difficulties (Sigafous, Arthur, & O’Reilly, 2003), such as persons with autism spectrum disorders (ASD) and/or intellectual disabilities (ID)
or developmental disabilities (DD\textsuperscript{1}; Emerson et al., 2001; Hill & Furniss, 2006; Matson, Dixon, & Matson, 2005). Indeed, 13–30\% of children and 88\% of adults with ASD (Horner, Carr, Strain, Todd, & Reed, 2002; McCarthy et al., 2010) and 10–15\% of individuals with ID/DD exhibit challenging behaviours (Emerson et al., 2001; Lowe et al., 2007). Challenging behaviours are associated with functional difficulties in adult life in the general population (see Moffitt, 1993; Robins, 1978; West & Farrington, 1977) and in ASD specifically (Bhaumik, Branford, McGrother, & Thorp, 1997; Marriage, Wolverton, & Marriage, 2009; Matson & Nebel-Schwalm, 2007; Matson & Rivet, 2008). These statistics speak to the importance of addressing challenging behaviours displayed by students with ASD and/or ID/DD enrolled in formal preschool programs and K–12 schools. Unfortunately, there are no reviews specific to this target population. As such, in the present review, we summarize the evidence regarding interventions for challenging behaviours among students (ages 3–21 years) with ASD and/or ID/DD that are suitable for school contexts.

Common categories of challenging behaviour include (a) physical and verbal aggression towards others (e.g., spitting, hitting, and biting) and property destruction; (b) disruptive behaviour (e.g., inappropriate verbalizations, temper tantrums); (c) stereotyped (repetitive) and self-stimulatory behaviour; and (d) self-injurious behaviour (e.g., self-biting, skin picking, head punching, and head hitting; Heyvaert, Maes, & Onghena, 2010; Matson & Nebel-Schwalm, 2007; Matson & Rivet, 2008; Matson, Terlonge, & González, 2006; Waters & Healy, 2012)\textsuperscript{2}. Related behaviours may include inadequate regard for boundaries and personal space. Unless these particular behaviours are considered intense, frequent, and/or long lasting and compromise physical safety and/or significantly threaten learning, they are generally not characterized as challenging behaviours. The impact of challenging behaviours is often pervasive in that the behaviours can occur across a variety of classroom activities (Chiang, 2008) and pose problems for the student with ASD and/or ID/DD and their teachers and peers. Indeed, of all the social and learning problems displayed by students with ASD and/or ID/DD, challenging behaviours are the most difficult for teachers (Bambara & Kern, 2005).

Job dissatisfaction may result when school staff members are not equipped with the knowledge and skills for dealing effectively with such behaviours. Some teachers identify challenging behaviours exhibited by all students as a primary reason for leaving the profession; in particular, school administrators report high rates of teacher attrition in the special education area and identify difficulty filling these specific positions (Ingersoll, 2001). Student well-being is compromised if efficient and effective response to these behaviours is delayed or absent. For example, students may repeatedly injure themselves, leading to infections and decreased physical health. Alternatively, classmates who feel threatened by a student’s aggressive behaviours will not feel safe, and as such, their mental health and learning may be compromised. Moreover, if children are frequently engaging in challenging behaviours, they may be distracted from learning opportunities, distract others from such opportunities, and be excluded or isolated by their peers, family members, or even school staff (Sprague & Rian, 1993). Conversely, intervention may prevent injuries and poor physical health, reduce secondary disabilities associated with the con-

\textsuperscript{1}The terms intellectual disability (ID) and developmental disabilities (DD) are sometimes used interchangeably in North America. In this paper, we used the term DD to include a wide range of neurological diagnoses that include, but are not limited to ID.

\textsuperscript{2}Sexualized behaviour is an additional challenging behaviour category; however, we did not find any studies in our literature search that addressed this type of behaviour, and as such, we have excluded the term from this review.
diation (including mental and social health), and improve the opportunities for learning and adaptive skill development.

Because of the debilitating and frequent nature of challenging behaviours, it is of utmost importance that they be appropriately assessed. Interestingly, in their 2007 review of the research on the assessment of challenging behaviours that commonly coexist in children with ASD, Matson and Nebel-Schwalm noted that research studies rarely include challenging behaviours (self-injury, aggression, and stereotypies, in particular) as an outcome variable. Related to this finding is a general lack of research on the best way to assess severe challenging behaviours in children with ASD, despite research that demonstrates the effectiveness of functional assessment. A third major problem is that challenging behaviours are generally not regarded as a core feature of ASD (with the exception of stereotypies), thus, strategies for intervention are not typically included as part of the diagnostic workup. Matson and Nebel-Schwalm also discussed that little work addresses differentiation of the various goals of assessment of challenging behaviours, such as measuring rate and intensity, general screening, identifying environmental factors and other predictor variables, and treatment and assessment outcomes.

Matson and Lovullo (2008) reviewed behavioural treatments for self-injurious behaviours in persons with ASD, finding that variants of behaviour-based treatments (e.g., choice-making, evaluating environmental factors, mand training, and replacement behaviours), punishment procedures, and pharmacotherapy seem most effective, although they also note many limitations and publication biases in this research. For example, the authors comment that studies describing cases where positive approaches (i.e., reinforcement or functional assessment) have failed to reduce self-injurious behaviour often include successful aversive add-ons (i.e., medication or punishment procedures). Importantly, however, studies with successful positives and failed aversive treatments, to our knowledge, have not been published. In their review of the literature on Applied Behaviour Analysis, Brosnan and Healy (2011) examined interventions employing antecedent manipulations, reinforcement, or consequences to reduce aggression in persons with ID/DD and found that all interventions were reported to be effective. Studies using functional assessment prior to choosing and administering the intervention reported the greatest reduction in aggressive behaviour—some studies reporting reductions to zero or near zero levels. Although not the focus of this review, it is important to note that there are systematic reviews of psychopharmacology for children with ASD (e.g., McPheeters et al., 2011), where medications (e.g., atypical antipsychotic agents, stimulants, and serotonin-reuptake inhibitors) to reduce interfering maladaptive behaviours have been reviewed (e.g., Myers & Johnson, 2007).

Horner et al. (2002) conducted a comprehensive synthesis of interventions for challenging behaviours in children with autism ages 8 years and younger, examining meta-analyses published from 1988 to 2000 and any original research published between 1996 and 2000. The reviewers applied an intervention effectiveness procedure (see Carr et al., 1999; see also the details of this procedure in Intervention Effectiveness subsection under Methods) to assist in ranking of strategies. Aggression/destruction, disruption, self-injury, and stereotypy were the behaviours most often targeted in published research. As is the case for many reviews, there was a large variation in the types of approaches examined. Studies dated 1990–2000 tended to employ stimulus-based or instruction-based interventions, while those prior to 1990 tended to be consequence-based (for definitions see Intervention Classification subsection under Methods). Early (prior to age 8 years) behaviour interventions reduced challenging behaviours by 80–90%, and interventions developed based on data from functional assessment appeared to be more effective, although there was a trend to use interventions not supported by data collected during the func-
tional assessment that was carried out prior to the administration of the intervention (Horner et al., 2002). Finally, Horner et al. concluded that no single intervention was uniquely effective for young children with autism.

Harvey, Boer, Meyer, and Evans (2009) reviewed interventions for challenging behaviours for individuals with ID/DD. The authors found that teaching replacement skills, combined with either systems change or antecedent/consequence manipulation, had the strongest effect and interventions preceded by functional analysis had larger effects than those that did not. Harvey et al. also found that different interventions seemed to work better for different behaviours. For example, changing antecedents effectively eliminated destructive and self-injurious behaviour, while changing consequences produced a medium effect in improving social behaviours. Skill replacement produced a medium effect across many challenging behaviours (stereotypic, destructive, self-injurious, and inappropriate social behaviours), while incorporating systems change with another intervention produced a moderate effect. Ultimately, Harvey et al. advocated for multi-component interventions that include functional analysis to inform skill replacement.

Heyvaert, Maes, Van den Noortgate, Kuppens, and Onghena (2012) conducted a meta-analysis of 30 studies describing pharmacological, psychotherapeutic, and contextual interventions (used alone or in combination) for challenging behaviours in persons with ID/DD. On average, most interventions had a large and statistically significant positive treatment effect; yet individual participant characteristics produced large variations in response to intervention. Heyvaert et al. noted that interventions incorporating the manipulation of antecedents had a larger effect and identified moderators (pre-treatment functional assessment, sensory impairment, study design, duration of treatment, diagnosis of autism, and gender) that affected intervention success.

The majority of published reviews document one-on-one interventions for challenging behaviour in a therapeutic or care facility. To our knowledge, there are only three reviews that explicitly summarize the research regarding interventions administered in formal preschool and K–12 school settings (i.e., Conroy, Dunlap, Clarke, & Alter, 2005; Machalicek, O’Reilly, Beretvas, Sigafoos, & Lancioni, 2007; Thompson, 2011). Conroy et al. (2005) focused their review on “proactive, positive strategies to prevent challenging behaviors” (p. 158) for young children (under age 6 years). ‘Positive strategies’ involve the use of environmental manipulation of antecedents, positive consequences, and instructional skill training to reduce challenging behaviours. Conroy et al. found that the research in the area was promising, but there is not yet sufficient research to claim that “positive in-school interventions” are evidence-based. Machalicek et al. (2007) reported that all categories of interventions for challenging behaviours in school settings that they examined (i.e., antecedent manipulation, changing instructional context, differential reinforcement, and self-management strategies) were effective at reducing the problematic behaviours. The reviewers indicated, however, that there are many methodological concerns or missing information in the studies they analyzed. In contrast to findings in other settings, there was no significant difference for those who included some sort of functional assessment procedure over those who did not (e.g., social stories, video modeling) in school settings. Thompson (2011) found a paucity of research specific to interventions for challenging behaviours in school settings. He advocated single-case research designs as an avenue to address the lack of evidence for interventions in school settings specifically.
Current Review

Although there are at least three reviews of challenging behaviour interventions in school settings (e.g., Conroy et al., 2005; Machalicek et al., 2007; Thompson, 2011), there are no syntheses of findings specific to individuals with ASD and/or ID/DD enrolled in formal preschool programs or K–12 schools. Thus, our review was targeted to characterize recent evidence regarding interventions for challenging behaviours among students enrolled in formal preschool programs and K–12 schools (i.e., students ages 3–21 years)\(^3\) with ASD and/or ID/DD in school settings, and to evaluate their effectiveness using established quality assessment (Tate et al., 2008) and intervention effectiveness criteria (Carr et al., 1999). Although the age range in our review differs, our goals and methodology were similar to Horner et al.’s (2002), but we limited our search to the years 2000–2011 inclusive.

Method

Search Strategy

We reviewed the literature pertaining to intervention strategies aimed at managing, decreasing, or eliminating challenging behaviour among individuals with ASD and/or ID/DD. We conducted an electronic search for studies published between January 2000 and November 2011 using the databases CINAHL, Cochrane Library, Educational Resources Information Center (ERIC), PsychINFO, PubMed, Social Work Abstracts, and Web of Science. Our search used the following keywords and descriptors with no limitations on the language of article: (a) mental retardation or developmental disability or cognitive disability or intellectual disability or autism or cognitive delay; and (b) misbehaviour* or spitting or acting out or aggression or aggressive or biting or tantrums or maladaptive behaviour* or behaviour disturbance* or self-injury* or stereotypy; and (c) school or student or education*; and (d) strategy or training or method* or intervention or treatment or program or instruction. These descriptors were used in singular and plural forms as well as in American and British spelling variations. Duplicate references were deleted using a pre-established database hierarchy. References were searched, screened, and stored through RefWorks 2.0 (RefWorks Cos, Bethesda, MD). In all, 414 references were located in this initial journal article search.

Selection of Studies for Targeted Review

The abstract and methodology sections of each of the 414 articles were reviewed to determine inclusion eligibility based on the criteria described below, which further limited the number of articles to review. All the articles we found and reviewed were written in English.

First, a peer-reviewed journal must have published the study. Book chapters, literature reviews, and dissertation studies were not included in this review. Second, to reflect the preschool and school population with whom our school team works, study participants must have been 3–21 years of age. Third, participants must have been identified as having a diagnosis of ASD (autism, autistic disorder, Asperger’s disorder, pervasive developmental disorder-not otherwise

---

\(^3\) In Manitoba, Canada (where this review took place), preschool to school age for students with developmental disabilities extends from 3 to 21 years of age. This is also true in other Canadian provinces and in the United States; as such, we reviewed research articles that reported interventions that were administered in preschool and school settings with students within this age range.
specified), mental retardation, or as having an ID or DD. Fourth, the study must have identified at least one of the challenging behaviours as a dependent variable. Fifth, the study must have reported original data about the effectiveness of at least one intervention aimed at managing, decreasing, or eliminating the challenging behaviour. Sixth, the study must have examined intervention strategies that were implemented by educators (i.e., teachers or educational assistants) in the school setting. Articles presenting interventions to be carried out exclusively by highly trained mental health professionals (e.g., psychologists, psychiatrists) and other individuals not associated with schools (i.e., parents and siblings) were excluded. Seventh, studies of interventions that were carried out in a non-school setting (e.g., hospital, therapist’s office) were excluded from review. Finally, studies that presented medical or alternative medicine treatments (e.g., medication and acupuncture) as the only method of intervention or in addition to behavioural treatment were not included in this review. We did not exclude studies that reported that participants were taking medication prior to intervention when the medication was not part of the intervention for the challenging behaviour (i.e., when the medication was being taken in baseline as well as treatment phases).

**Data Extraction**

Each study was analyzed across several categories: (a) research design; (b) participant characteristics, including number of participants, diagnosis, sex, and age; (c) description of intervention strategy; (d) description of dependent variables (i.e., targeted challenging behaviours); (e) location of intervention (e.g., classroom or another room within a school); (f) intervention delivery method (i.e., individual or group administration); (g) description of intervention effects on targeted challenging behaviours; and (h) confirmation of whether a functional assessment of behaviour was conducted prior to intervention delivery.

**Intervention Classification**

We classified the types of interventions used to reduce challenging behaviours and/or increase desired behaviours according to seven categories. The categories are non-exclusive in the sense that interventions described in our sample of articles typically consisted of multiple components that represented different categories.

1. **Stimulus-Based Procedures**: “Procedures that involved altering antecedent events prior to a problem behaviour” (Horner et al., 2002, p. 427) or desired behaviour that needs some intervention. Within this category, we noted the use of motivating operations (e.g., establishing or abolishing operation), stimulus fading, stimulus discrimination training, respondent/classical/Pavlovian conditioning, unfaded prompting, and other stimulus-based procedures not specified in advance.

2. **Instruction-Based Procedures**: Procedures that involved teaching appropriate behaviours that the student previously showed at zero or very low frequency. This included functional communication training, instruction on self-management skills, and other instruction-based procedures not specified in advance.

3. **Extinction-Based Procedures**: “Intervention procedures designed to withhold or minimize delivery of presumed reinforcers following problem behaviour” (Horner et al., 2002, p. 428).

4. **Reinforcement-Based Procedures**: “Procedures designed to increase or maintain desired behaviours through [response-dependent] delivery of events” (Horner et al., 2002, p. 428). We noted in this category the use of differential reinforcement of other (DRO) or
interventions for challenging behaviours, or differential reinforcement of low rate (DRL), token systems, non-contingent reinforcement, and other reinforcement-based procedures not specified in advance.

5. Punishment-Based Procedures: “Intervention procedures designed to reduce problem behaviour via delivery of contingent events” (Horner et al., 2002, p. 428). We included in this category time-outs, overcorrection, reprimands, response cost, and other punishment-based procedures not specified in advance.

6. Systems Change: “Intervention procedures that were designed to alter structural features of an environment...including change in staffing pattern, change in outcome measures used to assess success, and change in administrative oversight of the intervention implementation” (Horner et al., 2002, p. 428).

7. Unclear/Other: Interventions that were not described with sufficient detail to determine the operating principles or did not fit any of the other categories.

Interventions that included Functional Communication Training or self-management were classified as Instruction-Based Procedures rather than being decomposed into component parts (e.g., Reinforcement-Based, Stimulus-Based). We did this in order to limit redundancies with other categories and to avoid eliminating the intended designation of Instruction-Based Procedure. As mentioned above, procedures that included direct instruction on appropriate behaviours that the student initially showed at zero or very low frequency were classified as Instruction-Based Procedures.

While reviewing each article, we noted the use of the broad term functional assessment—sets of instructions for defining behaviour, identifying the stimuli that evoke and/or maintain the behaviour, and using this information to determine appropriate interventions (e.g., Stichter, 2001). Several types of functional assessment exist including those involving indirect assessment (based on interviews and surveys); direct observation of the antecedents-behaviour-consequence relations in the natural environment; and experimental functional analysis, in which environmental conditions are manipulated in order to observe their effect on the frequency of the behaviour. If sufficient detail was provided to indicate the type of functional assessments conducted, it was recorded as a functional analysis, (e.g., Iwata, Dorsey, Slifer, Bauman, & Richman, 1982), and/or a descriptive analysis (e.g., direct observations, parent or teacher interview).

One or two observers scored each article as either including or not including 29 discrete intervention categories. Observer A was the primary rater for 21 articles; Observer B was the primary rater for 17 articles. Each observer also classified six of the other primary rater’s articles to provide inter-observer reliability checks on 31% (12/39) of the sample. Most articles included only a small fraction of possible categories (M = 19%); thus, inter-observer agreement was determined using Cohen’s Kappa to account for chance agreement on occurrence and non-occurrence of categories. Kappa values on the 12 articles ranged from 0.71–1.00 (M = 0.91).

**Intervention Effectiveness**

Intervention effectiveness was calculated for interventions used to reduce problem behaviours. The absence of group-based experimental designs, in general, and randomized controlled-trials, in particular, precluded the use of standard meta-analytical procedures. Instead, we used a percent reduction measure (Carr et al., 1999) to calculate intervention effectiveness. Data points were measured from article graphs to determine the value of the last three data points of the
baseline phase and the last three data points of the intervention phase. Percent reduction was then calculated as the difference between the mean of the last three baseline data points and the mean of the last three intervention data points, divided by the mean of the last three baseline data points, multiplied by 100. Intervention effectiveness was measured for all available graphs (i.e., for each participant and/or for each intervention procedure). In the case of reversal designs (e.g., ABAB), we used the guidelines outlined in Carr et al. (1999). That is, the last three data points of the first baseline phase were compared to the last three data points of the last intervention phase.

Percent reduction of problem behaviour was averaged across participants within studies or within each intervention if the study reported the effects of two or more distinct interventions for comparison purposes (i.e., where these interventions were designed not as minimal- or no-treatment control conditions, but as treatments that could plausibly be presented independently). The authors of two articles (Anderson & Le, 2011; Devlin, Healy, Leader, & Hughes, 2010) provided percent reduction scores for each of the two interventions that they examined.

**Quality Assessment**

To assess critically the methodological quality of the studies included in the present review, the researchers used the Single-Case Experimental Design (SCED) Scale (PsycBITE, 2010). The PsycBITE™ Committee developed the SCED Scale to provide a means for rating the methodological quality of studies employing single-subject designs. Tate et al. (2008) provided information on the development, reliability, and content validity of this scale. Given the dominance of single-subject designs in the behavioural intervention literature, the SCED Scale was determined to be the most appropriate quality assessment rating scale for our purposes. The scale consists of 10 items pertaining to methodological quality (e.g., experimental design, reliability assessment, and generalization) and use of statistical analysis and 1 item pertaining to participant demographics. Studies receive a score of one for each item that clearly meets criteria outlined in the SCED Scale raters’ manual (PsycBITE, 2010) providing an overall methodological quality score, with a maximum of 10.

The SCED Scale raters’ manual (PsycBITE, 2010) specified instructions for scoring reversal and multiple-baseline designs, but lacked instructions for alternating treatments designs (ATDs). The maximum potential rating was 10 for multiple baseline and A-B-A reversal designs (as per SCED guidelines), 9 for variations of reversal designs (e.g., B-A-B, A-B-C-A), and 8 for A-B designs. Because the manual did not explicitly state that ATDs be excluded, and given that ATDs are often used in single-subject experiments (Kazdin, 2010), we contacted the primary author regarding inclusion of these designs. The author affirmed that ATDs were admitted for the published validation study of the SCED (R. Tate, personal communication, June 27, 2012), so we included articles that used ATDs into our assessment without restrictions. That is, articles that reported an ATD were eligible to earn points for SCED items 3, 5, and 6, as well as item 4, if a baseline phase occurred prior to the alternating treatments phase (Tate et al., 2008).

Five independent raters evaluated studies included in the review with overlap on 25% of studies in order to provide inter-rater agreement scores. Ratings were compared on an item-by-item basis, and an agreement score was computed for each article by dividing the number of agreed-upon items by the total number of items (10) and expressing the result as a percentage. The mean inter-rater agreement was 92% (range: 80–100%). Discrepancies were resolved by consensus. No studies were excluded using this method.
Results

Article Characteristics

Thirty-eight articles met our inclusion criteria. The studies reviewed were based primarily in the United States (35 studies), with one article each based in the United Kingdom, Republic of Ireland, and Australia. The articles were mainly extracted from journals focusing on behaviour interventions (26 studies) and autism and/or ID/DD (6 studies) research. The remaining articles were obtained from journals focusing on social work (1 study), therapy (2 studies), and education (3 studies). The number of articles per year pertaining to challenging behaviours remained consistent over time (range: 1–4 studies), with the exception of 2005 and 2011, in which seven and nine articles, respectively, pertaining to challenging behaviours were published.

Participant Demographics

Our targeted review identified 68 participants (56 males, 12 females) across the 38 studies that met inclusion criteria. These participants ranged from 3–21 years of age ($M = 8.4$ years, $SD = 3.8$). Two studies, representing five participants, did not specify their participants’ ages, but did report either age ranges of the students (i.e., Johnson, Van Laarhoven, & Repp, 2002) or indicated that the students were elementary school-aged (i.e., Glaeser, Pierson, & Fritschmann, 2003). The majority of participants (71.4%) reportedly had a primary diagnosis of ASD. The remaining participants had primary diagnoses of ID (15.7%), Asperger’s Disorder (4.3%), Pervasive Developmental Disorder Not Otherwise Specified (4.3%), and other developmental disorders (4.3%). In Table 1, we present a summary of participant characteristics. It is important to note that most authors did not provide sufficient information about the ability level or comorbidities of the students in the study to enable classification or comment.

Research Approaches and Designs

Of the 38 articles reviewed, 33 involved quantitative approaches, 1 was qualitative, and 4 used mixed methods. Of the quantitative studies, 27 used some form of single-subject experimental design: 10 reported a multiple baseline design, 10 used a reversal design, 6 used a multi-element design, and 1 used a combination of two or more single-subject designs. Three articles reported quantitative non-experimental case studies, and three reported other types of designs.

Behaviours Targeted

Four broad categories of challenging behaviours (see Introduction) were targeted by intervention in the identified literature. In the 38 studies that we reviewed, stereotyped, repetitive, and self-stimulatory behaviour was the most commonly targeted challenging behaviour (20 studies), followed by physical and verbal aggression and property destruction (12 studies), self-injurious behaviour (11 studies), and disruption (8 studies). It is important to note that studies may have reported effects on multiple categories of challenging behaviours. In all, 68% of studies reported measuring effects on a single category of challenging behaviour, while the remaining 32% reported effects on multiple categories (see Table 1).

Intervention Classification

The majority of articles (30) reported the use of some type of functional assessment. Of these, 17 studies employed functional analysis, 16 employed descriptive analyses, 5 employed
Table 1
Characteristics of 38 Reviewed Articles Organized by Category of Challenging Behaviours

<table>
<thead>
<tr>
<th>Challenging Behaviour Category Reference</th>
<th>Participant Characteristics</th>
<th>Intervention(s) Category (Specifics)</th>
<th>Quality Assessment</th>
<th>Intervention Effectiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Physical and verbal aggression towards others and property destruction</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cannella-Malone, Tullis, &amp; Kazee (2011)</td>
<td>3 males, ASD, ID/DD, Emotional Behaviour Disorder</td>
<td>Stimulus-based; Reinforcement-based (Exercise)</td>
<td>8</td>
<td>87%</td>
</tr>
<tr>
<td>Frea, Arnold, &amp; Vittimberga (2001)</td>
<td>1 male, ASD, ID/DD</td>
<td>Instruction-based (Picture exchange)</td>
<td>7</td>
<td>100%</td>
</tr>
<tr>
<td>Buggey (2005)</td>
<td>5 males, ASD</td>
<td>Stimulus-based (Video self-modeling)</td>
<td>7</td>
<td>94%</td>
</tr>
<tr>
<td>Buckley &amp; Newchok (2005)</td>
<td>1 male, ASD</td>
<td>Instruction-based (Functional communication training)</td>
<td>5</td>
<td>81%</td>
</tr>
<tr>
<td>Foxx &amp; Meindl (2007)</td>
<td>1 male, ASD, Disruptive Behaviour Not Otherwise Specified</td>
<td>Stimulus-based*</td>
<td>4</td>
<td>100%</td>
</tr>
<tr>
<td>Ramsey, Waller, &amp; Walker (2005)</td>
<td>1 male, ID/DD</td>
<td>Stimulus-based; Reinforcement-based (Visual cues*)</td>
<td>4</td>
<td>63%</td>
</tr>
<tr>
<td><strong>Physical and verbal aggression towards others and property destruction; Stereotypy, repetitive, and self-stimulatory behaviour</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stichter, Randolph, Kay, &amp; Gage (2009)</td>
<td>3 males, ASD</td>
<td>Stimulus-based; Extinction-based; Unclear/other (Antecedent-based)</td>
<td>7</td>
<td>79%</td>
</tr>
<tr>
<td>Bartlett, Rapp, Krueger, &amp; Henrickson (2011)</td>
<td>1 male, ASD</td>
<td>Stimulus-based; Reinforcement-based; Punishment-based (Response cost)</td>
<td>6</td>
<td>98%</td>
</tr>
<tr>
<td>Luiselli, Ricciardi, Schmidt, &amp; Tarr (2004)</td>
<td>1 male, ASD</td>
<td>Reinforcement-based (DRO)</td>
<td>5</td>
<td>100%</td>
</tr>
<tr>
<td><strong>Physical and verbal aggression towards others and property destruction; Self-injurious behaviour</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Van Camp, Lerman, Kelley, Contrucci, &amp; Vorndran (2000)</td>
<td>1 male, 1 female, ID/DD</td>
<td>Extinction-based; Reinforcement-based (Noncontingent reinforcement)</td>
<td>8</td>
<td>98%</td>
</tr>
<tr>
<td>Hammond, Iwata, Fritz, &amp; Dempsey (2011)</td>
<td>3 males, 1 female, ASD, ID/DD, LD, Seizure Disorder</td>
<td>Reinforcement-based (DRO)</td>
<td>7</td>
<td>96%</td>
</tr>
<tr>
<td>Braithwaite &amp; Richdale (2000)</td>
<td>1 male, ASD, ID/DD</td>
<td>Stimulus-based; Instruction-based; Reinforcement-based (Functional communication training)</td>
<td>6</td>
<td>100%</td>
</tr>
<tr>
<td>Foxx &amp; Garito (2007)</td>
<td>1 male, ADHD, ASD, ID/DD</td>
<td>Stimulus-based; Punishment*</td>
<td>6</td>
<td>99%</td>
</tr>
<tr>
<td>O’Connor Sorenson-Burnworth, Rush, &amp; Eidman (2003)</td>
<td>1 male, ADHD, Fragile X Syndrome, ID/DD</td>
<td>Stimulus-based; Reinforcement-based; Punishment-based*</td>
<td>6</td>
<td>98%</td>
</tr>
<tr>
<td>Challenging Behaviour Category Reference</td>
<td>Participant Characteristics</td>
<td>Challenging Behaviour</td>
<td>Reference</td>
<td>N and Gender</td>
</tr>
<tr>
<td>-----------------------------------------</td>
<td>----------------------------</td>
<td>-----------------------</td>
<td>-----------</td>
<td>--------------</td>
</tr>
<tr>
<td>Physical and verbal aggression towards others and property destruction; Self-injurious behaviour (Cont’d)</td>
<td>Devlin et al. (2010)</td>
<td>4 males</td>
<td>ASD, Epilepsy, LD</td>
<td>6–11</td>
</tr>
<tr>
<td>Disruption</td>
<td>Galiatsatos &amp; Graff (2003)</td>
<td>1 male</td>
<td>ASD, OCD</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>Glaeser et al. (2003)</td>
<td>1 female</td>
<td>ASD</td>
<td>not specified</td>
</tr>
<tr>
<td>Disruption; Stereotyped, repetitive, and self-stimulatory behaviour</td>
<td>Hines &amp; Simonsen (2008)</td>
<td>1 male</td>
<td>ASD</td>
<td>3</td>
</tr>
<tr>
<td>Disruption</td>
<td>Soares, Vannest, &amp; Harrison (2009)</td>
<td>1 male</td>
<td>ASD</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>Cale et al. (2009)</td>
<td>1 male, 2 females</td>
<td>ASD</td>
<td>5–8</td>
</tr>
<tr>
<td>Stereotyped, repetitive, and self-stimulatory behaviour</td>
<td>Mancina, Tankersley, Kamps, Kravits, &amp; Parrett (2000)</td>
<td>1 female</td>
<td>ASD</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Ahrens, Lerman, Kodak, Worsdell, &amp; Keegan (2011)</td>
<td>4 males</td>
<td>ASD</td>
<td>4–6</td>
</tr>
<tr>
<td></td>
<td>Nuzzolo-Gomez, Leonard, Ortiz, Rivera, &amp; Greer (2002)</td>
<td>3 males, 1 female</td>
<td>ASD</td>
<td>3–7</td>
</tr>
<tr>
<td></td>
<td>Haley, Heick, &amp; Luiselli (2010)</td>
<td>1 male</td>
<td>ASD</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Johnson et al. (2002)</td>
<td>2 males, 2 females</td>
<td>ID/DD</td>
<td>6–12</td>
</tr>
<tr>
<td></td>
<td>Liu-Gitz &amp; Banda (2010)</td>
<td>1 male</td>
<td>ASD</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Schumacher &amp; Rapp (2011)</td>
<td>1 male, 1 female</td>
<td>ASD</td>
<td>8, 5</td>
</tr>
<tr>
<td>Challenging Behaviour Category Reference</td>
<td>Participant Characteristics</td>
<td>干预(s) Category (Specs)</td>
<td>Quality Assessment</td>
<td>Intervention Effectiveness</td>
</tr>
<tr>
<td>-----------------------------------------</td>
<td>------------------------------</td>
<td>-------------------------</td>
<td>--------------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td>Stereotyped, repetitive, and self-stimulatory behaviour (Cont’d)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conroy et al. (2005)</td>
<td>1 male</td>
<td>ASD</td>
<td>6</td>
<td>Stimulus-based (Visual cues)</td>
</tr>
<tr>
<td>O’Connor, Prieto, Hoffmann, DeQuinzio, &amp; Taylor (2011)</td>
<td>1 male</td>
<td>ASD</td>
<td>11</td>
<td>Stimulus-based; Reinforcement-based; Punishment-based (Visual cues)</td>
</tr>
<tr>
<td>Anderson &amp; Le (2011)</td>
<td>1 male</td>
<td>ASD</td>
<td>7</td>
<td>Reinforcement-based, punishment-based: (Positive practice overcorrection; DRA)</td>
</tr>
<tr>
<td>Meador, Derby, McLaughlin, Barretto, &amp; Weber (2007)</td>
<td>1 female</td>
<td>Rett Syndrome</td>
<td>6</td>
<td>Reinforcement-based (DRO)</td>
</tr>
<tr>
<td>Longano &amp; Greer (2006)</td>
<td>3 males</td>
<td>ASD</td>
<td>5–7</td>
<td>Reinforcement-based (Stimulus-stimulus pairing)</td>
</tr>
<tr>
<td>Taylor, Hoch, &amp; Weissman (2005)</td>
<td>1 female</td>
<td>ASD</td>
<td>4</td>
<td>Stimulus-based; Reinforcement-based (DRO)</td>
</tr>
<tr>
<td>Groskreutz, Groskreutz, &amp; Higbee (2011)</td>
<td>1 male</td>
<td>ASD</td>
<td>4</td>
<td>Reinforcement-based (Response competition)</td>
</tr>
<tr>
<td>Kellet (2003)</td>
<td>1 male</td>
<td>ID/DD</td>
<td>8</td>
<td>Reinforcement-based; Unclear/other (Intensive interaction)</td>
</tr>
<tr>
<td>Cassella et al. (2011)</td>
<td>2 males</td>
<td>ASD</td>
<td>4–7</td>
<td>Stimulus-based; Unclear/other (RIRD)</td>
</tr>
<tr>
<td>Self-injurious behaviour</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O’Reilly, Sigafoos, Lancioni, Edrisinha, &amp; Andrews (2005)</td>
<td>1 male</td>
<td>ASD</td>
<td>12</td>
<td>Stimulus-based (Individualized activity schedule)</td>
</tr>
<tr>
<td>Carey &amp; Halle (2002)</td>
<td>1 male</td>
<td>ID/DD</td>
<td>12</td>
<td>Stimulus-based; Reinforcement-based (Contingent access to reinforcers)</td>
</tr>
</tbody>
</table>


*Methodological quality score based on the SCED scale (Tate et al., 2008). Scores range from 0 (lowest quality) to 10 (highest quality).

*Percent reduction in challenging behaviour. *Study authors indicated that the participant(s) were on medication but that the medication was not part of the intervention for the challenging behaviour.
both descriptive and functional analyses, and 2 studies did not specify the type of functional assessment. Each study reported 1–3 major intervention categories (M = 1.8, Mode = 1), and 1–7 procedures (M = 2.6, Mode = 2). Reinforcement-based interventions were the most common (23 articles), closely followed by stimulus-based interventions (22 articles). Punishment-based interventions were reported in 8 articles; response cost was used in 6 of these articles. None of the articles involved systems change.

Differential reinforcement (DRO, DRL, or DRI) was the most common reinforcement-based procedure (11 articles). Token systems and non-contingent reinforcement were used in 6 and 7 articles, respectively. Many reinforcement-based procedures were classified as “other,” meaning that the procedure did not fit into a pre-specified type of reinforcement-based procedure. Of the 8 “other” reinforcement-based procedures, altering (typically thinning) reinforcement schedules was seen in 7 studies. The remaining “other” reinforcement-based procedures involved establishing stimuli as conditioned reinforcers. In the stimulus-based category, unfaded prompting was the most frequent procedure (13 articles), followed by stimulus discrimination training (5 articles), and stimulus fading (5 articles).

Interventions that fit the Unclear/Other interventions category were recorded in 8/38 articles. Response interruption and redirection (RIRD) was evaluated in 6 of these 8 articles and were included in this category because clear evidence about the behavioural function(s) of this procedure, especially the interruption component, is still quite limited. Studies that described positive reinforcement (e.g., praise, social attention) for compliance with redirection were also coded as reinforcement-based. Additional Unclear/Other interventions were scored in that category because it was difficult to identify or analyze the function-based procedures they comprised (e.g., sensory integration therapy).

**Intervention Effectiveness**

Intervention effectiveness (i.e., reduction or substantial decrease in the problem behaviour) was calculated for 36 articles (see Table 1). Intervention effectiveness could not be calculated for qualitative articles (i.e., Glaeser et al., 2003) or those that did not report the frequency of problem behaviours (i.e., Cale, Carr, Blakeley-Smith, & Owen-DeSchryver, 2009). Two articles (i.e., Anderson & Le, 2011; Devlin et al., 2010) compared different intervention procedures (e.g., behavioural and sensory integration); thus, intervention effectiveness was calculated separately for each procedure. In total, 38 intervention effectiveness scores were calculated. Reduction of problem behaviour ranged from 100% to -12%; only a single intervention produced negative reduction (i.e., an increase in problem behaviour rates). Mean reduction was 81%, and scores were negatively skewed: 16/38 interventions produced 90% or greater reduction, 8/38 interventions produced 80–89% reduction, 4/38 interventions produced 70–79% reduction, 6/38 interventions produced 50–70% reduction, and 4 studies reported less than 50% problem behaviour reduction.

Inclusion or exclusion of the major intervention categories and most common procedures generally did not correlate with intervention effectiveness scores. Functional assessment (whether descriptive or analytic) was the intervention element that related most strongly with effectiveness. In the 30 studies that employed functional assessment, mean effectiveness was 85%. In the 8 studies that did not use functional assessment, mean effectiveness was 66%.

**Quality Assessment**

Quality scores were assessed for 37 of 38 articles (see Table 1); one article (Glaeser et al., 2003) was not rated using the SCED Scale’s inclusion criteria because it reported neither pre-
nor post-treatment data. Quality scores for the 37 remaining studies ranged from 0–8 ($M = 5.61$, $SD = 1.85$). More specifically, 15 studies obtained SCED scores equal to 7 or greater, 20 articles obtained scores that fell in the 4–6 range, and 2 studies obtained a SCED score equal to 3. Only one study (i.e., Cassella, Sidener, Sidener, & Progar, 2011) obtained a quality score of zero, and no studies obtained a perfect score of 10.

**Discussion**

Our goal for the present review was to identify interventions for reducing challenging behaviours displayed by students with ASD and/or ID/DD appropriate to the school context that showed evidence of effectiveness and strong support based on the quality of the research design. Consistent with previous studies (Horner et al., 2002; Heyvaert et al., 2012), we found few studies that examined any one specific intervention consisting of only a single procedure for any one specific target behaviour. Most studies reported a combination of procedures from the stimulus-based and reinforcement-based intervention categories. As such, it is difficult to determine precisely which components of an intervention are most essential and impactful in treatment (see Future Research). The majority of studies also reported the use of functional assessment procedures (see Limitations). Similar to findings reported by Conroy et al. (2005) and Thompson (2011), a strong emerging trend toward positive in-school interventions was evident in the last decade.

Stereotyped, repetitive, and self-stimulatory behaviours remain the most-studied category of challenging behaviours, and stimulus- and reinforcement-based interventions to treat them are commonly researched. Stimulus- and reinforcement-based strategies produced large percent reductions in challenging behaviour (i.e., averaging over 70% and 80% reductions, respectively), to the extent that it is difficult to say that one or the other type of intervention is especially effective for our population of interest. In fact, most interventions (24/38 studies) produced large (80–100%) reductions in challenging behaviour in nearly every category of such behaviour. These findings are consistent with the reductions in challenging behaviour reported in previous reviews (Brosnan & Healy, 2011; Heyvaert et al., 2012; Horner et al., 2002; Machalicek et al., 2007). Given the heterogeneity of interventions observed in the articles we reviewed, as well as variability in how often each intervention type occurred, there is little basis to generalize about the effectiveness of interventions relative to one another. However, similar to other reviews (Brosnan & Healy, 2011; Harvey et al., 2009; Horner et al., 2002), we found that interventions that included a functional assessment were more effective than interventions without (82% vs. 66%).

It is important to consider whether the percent reduction scores obtained in the studies corresponded to clinical significance. The studies we reviewed rarely addressed social validity (Kazdin, 1977; Wolf, 1978), making it impossible to answer this question based on our data. Others have reported a similar finding (Horner et al., 2002). It is not enough to quantify the level of improvement in the target behaviour, but it is also essential to evaluate whether this improvement is clinically and/or socially important for the student, peers, and teachers (see Future Research).

**Limitations**

Assessment of methodological rigour is not often conducted in review studies. Our goal was not only to summarize what studies have found, but also to communicate the degree by which these findings were reliable and appropriate to school contexts. Consequently, it was essential to evaluate the rigour of studies in the present review. Unfortunately, the SCED Scale
raters’ manual (PsycBITE, 2010) does not include a discussion on the interpretation of the quality score. It is unknown whether a score of 5, for example, indicates that a study showed acceptable rigor. Tate et al. (2008) report that on 85 articles assessed using their scale, the distribution of scores was normal with a mean score of 4.65 (SD = 2.35). Our observed mean of 5.61 (SD = 1.85) is comparable to Tate’s report. Scores of 7 or higher (15 articles in our sample) may therefore be considered to indicate equal or higher methodological quality than approximately two-thirds of single-case reports. Caution must be observed regarding simple SCED Scale totals, however, because users may wish to weight items differently for different quality assessment purposes. A complete report would require listing scores for all items. The next edition of the SCED will include an algorithm for producing cut-off scores (Tate, personal communication, June 27, 2012).

Another limitation should be noted. Our search consisted only of an electronic database search, and did not use ancestry searches (i.e., references cited in identified studies), nor did we conduct hand searches of relevant journals. Consequently, the studies we found are limited to those that were identifiable by keyword searches (as outlined in our methodology). Moreover, as noted earlier, publication bias (studies that do not produce a result are rarely published) inhibits our ability to comment on forms of intervention that have not successfully improved challenging behaviours.

Future Research

Based on our review, we noted several avenues for future research. First, determining which intervention components are indeed effective is often difficult, as the studies we reviewed often reported a combination of procedures from the stimulus-based and reinforcement-based intervention categories. Consequently, future research should continue to isolate aspects of programming to help determine effectiveness of individual components. Another important area for future research relates to whether improvements in behaviour are clinically and/or socially valid. While intuitively it seems that improvements in behaviour will likely parallel improvements in overall quality of life and opportunities for learning, it would be helpful to directly evaluate whether reductions in problem behaviours actually improve functioning and or learning for those who receive intervention, or improve the experiences of those working with people with ASD and/or ID/DDD. Finally, understanding whether reductions in challenging behaviour are significant to families (in the home or school context) is vital. While the focus of our review was specific to interventions conducted in school settings by school staff, it is likely that families who perceive significant improvement in behaviour may be more apt to implement similar procedures in home settings, which may in turn, improve overall intervention effectiveness.

Recommendations and Conclusions

Many reviews of interventions for challenging behaviours have been conducted over the past several decades; many are not specific to ASD or ID/DD and preschool and school settings. In order to translate research for teachers, it was essential that we identified interventions that could be applied in school settings by teachers and educational support staff. From this review, we conclude that school staff has various options available to intervene with challenging behaviours that are supported by evidence. The findings of the present review reinforce previous findings indicating that regardless of the intervention used, significant reduction in challenging behaviour occurs for most of the commonly referenced approaches. Use of functional assessment
procedures was associated with a stronger magnitude of impact, and when appropriate, should be a starting point for interventions.

Consequently, the practical implications of this study are that teachers have a variety of options for intervening with challenging behaviours that are likely to be effective. If school teams aim to use evidence-based practices, it is best to begin with procedures that have documented effectiveness before moving to novel approaches to changing behaviour. For interventions based in behavioural psychology, the evidence supports the use of functional assessments prior to intervening. However, for approaches that may not be primarily behavioural (e.g., social stories), a functional assessment may not impact the intervention success (Machalicek et al., 2007). We recommend that teachers measure the effect of an intervention, regardless of the type chosen, as this is the most useful way to evaluate whether an intervention actually reduces the occurrence of challenging behaviour (Kratochwill & Shemoff, 2003).

In a practical sense, it is important that researchers not only communicate research findings to school teams, but that they do so in a way that is easily consumed by the individuals intended to apply them. As such, it is essential that we provide tools to teachers that can demystify or clarify research so that it is easily understood and applied to real life school contexts. Our research group is engaged in projects to make these research findings more concrete by creating fact sheets, scripts, and decision trees school staff can use to make decisions about which interventions to use (see www.katydid.info). In this way, we hope to improve the educational experiences of individuals with ASD and/or ID/DD and their peers. Ultimately, we hope this will impact school staff’s ability to intervene with students and improve the educational experience for all staff and students.

References

References marked with an asterisk indicate studies included in the review.


Interventions for Challenging Behaviours


Montgomery et al.


Interventions for Challenging Behaviours


Authors’ Note

Correspondence concerning this article should be addressed to Janine Montgomery, Department of Psychology, University of Manitoba, Winnipeg, MB, Canada, R3T 2N2.

Email: Janine.Montgomery@umanitoba.ca

This project was funded by the Canadian Institute of Health Research. We wish to thank Drs. J. Virues Ortega and C. T. Yu and two anonymous reviewers for their reviews and suggestions for this paper.