The Journal of Child Psychology and Psychiatry

Journal of Child Psychology and Psychiatry **:* (2025), pp **–** doi:10.1111/jcpp.14177

Research Review: Conceptualizing and measuring 'problem behavior' in early intervention autism research – a project AIM secondary systematic review

Kristen Bottema-Beutel,[1] Ruoxi Guo,[1] Jessica Hinson-Wiliams,[1] Yueyang Shen,[1] Shannon LaPoint,[2] Tiffany Woynaroski,[3,4,5,6,7] and Micheal Sandbank[8]

[1]Lynch School of Education and Human Development, Boston College, Boston, MA, USA; [2]School of Teacher Education, Florida State University, Tallahassee, FL, USA; [3]Department of Hearing and Speech Sciences, Vanderbilt University Medical Center, Nashville, TN, USA; [4]Frist Center for Autism and Innovation, Vanderbilt University, Nashville, TN, USA; [5]Vanderbilt Brain Institute, Vanderbilt University, Nashville, TN, USA; [6]Vanderbilt Kennedy Center, Vanderbilt University Medical Center, Nashville, TN, USA; [7]Department of Communication Sciences and Disorders, John A. Burns School of Medicine, University of Hawaii at Manoa, Honolulu, HI, USA; [8]Division of Occupational Science and Occupational Therapy, Department of Health Sciences, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

Background: Some autistic children exhibit behavior that caregivers, clinicians, and researchers consider problematic. However, there is little consensus about the types of behaviors that should be treated as a problem and reduced via intervention. In autism intervention research, problem behaviors range from inherently harmful behaviors such as aggression and self-injury to nonnormative but not harmful behaviors associated with autism such as repetitive movements. Likewise, there are a variety of conceptualizations and measurement practices used to assess these behaviors. **Methods:** In this secondary systematic review of group-design, nonpharmacological intervention studies for autistic children up to age eight, we explore researchers' conceptualizations of problem behavior and measurement systems to assess problem behavior. We defined problem behavior as any outcome where behaviors were targeted for reduction or elimination. A coding scheme was applied to 102 studies that met inclusion criteria for the secondary review. All studies were double coded by two independent coders. **Results:** Sixty-two percent of studies described reducing behavior as a primary or secondary purpose of the study and/or intervention, 33% gave a rationale for targeting behaviors for reduction, and 28% offered a conceptualization of the behavior(s) they targeted. Only 8% offered a conceptual definition. The most common measures were 'off-the-shelf' measures that had undergone at least some previous validation beyond interrater reliability and that involved parent reports. For the 10 most common assessment measures, two were validated along six different validation dimensions in autistic populations. All but one full scale or subscale measured behaviors that were nonnormative but not inherently harmful, or a mix of behaviors that were inherently harmful and that were nonnormative but not inherently harmful. **Conclusions:** Intervention researchers should provide clear definitions and rationales for targeting behaviors for reduction via intervention and should develop refined measurement tools for assessing these behaviors in collaboration with the autistic community. **Keywords:** Problem behavior; restrictive and repetitive behavior; autism; early intervention; treatment; measurement.

Introduction

Some parents of young autistic children report that aspects of their children's behavioral repertoires negatively impact their autistic children and their families (Leadbitter et al., 2018). For example, autistic preschoolers are more likely to act aggressively (e.g., hitting, biting) toward others as compared to nonautistic preschoolers, and this can impede the development of autistic children's peer relationships and decrease family quality of life (Sivathasan et al., 2024). Behaviors considered undesirable are often referred to in autism literature as 'problem behaviors', and reducing their occurrence is a common early childhood intervention target for young autistic children (Sandbank et al., 2023). However, there is little agreement among researchers about the kinds of behaviors that should be considered a problem, the root causes of these behaviors, and the best approach to supporting autistic children who exhibit such behavior.

The current study is a secondary systematic review of a previously conducted meta-analysis of nonpharmacological intervention research conducted with autistic children up to age eight (Sandbank et al., 2023). In this secondary analysis, we examined studies that included an outcome in which a behavior (e.g., self-injury) or group of behaviors (e.g., restricted and repetitive behaviors) were targeted for reduction, to determine how autism intervention researchers conceptualize, address, and measure behavior considered problematic. In this study, we use 'problem behavior' to refer to the collection of behaviors that intervention researchers have

Conflict of interest statement: See Acknowledgements for full disclosures.

© 2025 The Author(s). *Journal of Child Psychology and Psychiatry* published by John Wiley & Sons Ltd on behalf of Association for Child and Adolescent Mental Health.
This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

attempted to decrease or eliminate in autistic children, because this is a term commonly used in the literature. However, we note this term is problematic and potentially stigmatizing. In our discussion, we suggest alternative approaches to defining and referring to this group of behaviors.

Defining and characterizing problem behavior

The considerable variation in the kinds of behaviors considered a problem is likely attributable to the fact that many autism researchers conceptualize problem behavior as comprising any behavior considered problematic for the people in proximity to those exhibiting it. For example, in a well-cited study aimed at estimating the prevalence of problem behavior among adults with intellectual disability receiving residential services (including those with co-occurring autism; Emerson, 2001), the authors considered problem behavior to be any behavior support staff considered to present 'a serious management problem or would do were it not controlled within that setting' (p. 79). Other influential studies offer similar broad definitions but reference the violation of larger social norms. Hanley, Iwata, and McCord (2003) considers any behavior socially significant enough that it causes others to complain about it to be a problem behavior. However, this definition does not elaborate on who constitutes 'others' or the nature of such complaints. What others may consider a problem is dependent on social context, understandings about autism and autistic behavior, and value judgments regarding norm violations.

There are likely some behaviors that researchers, autism community members, and people who belong to both groups may agree are problematic regardless of context. This could include behaviors with high potential for harm such as self-injury, aggression, and elopement. However, inadequate conceptualizations of what constitutes problem behavior in autism research, including little consideration of the multiple perspectives from which the behavior could be considered a problem or not, means *any* behaviors exhibited by autistic people could be subsumed under the category of problem behavior and subsequently targeted for reduction. This has been shown to be the case in an analysis of problem behavior outcomes in intervention research with transition-age autistic youth (Bottema-Beutel, McKinnon, Mohiuddin, LaPoint, & Kim, 2024), which found researchers often operationalized problem behavior in such a way that seemingly innocuous behaviors, such as finger tapping, looking at classmates, and the production of nonspeech vocalizations, were combined in the same general construct along with behaviors with high potential for harm, such as aggression and self-injury. Bottema-Beutel et al. (2024) also found authors of most intervention studies conceptualized problem behavior by providing lists of examples of such behavior, but they did not provide a definition that would allow readers to evaluate the extent to which any specific instance of a behavior meets definitional criteria of 'problem behavior' Again, the lack of a clear definition can make it difficult for researchers, clinicians, and families to clearly communicate about which behaviors present real problems, which behaviors could potentially lead to problems down the road, and which behaviors are simply part of being autistic and are not inherently problematic (see Carr, 2007; Kapp et al., 2019; Kim & Bottema-Beutel, 2019 for discussions).

Historically, researchers have made assumptions about the negative effects of autistic behaviors now subsumed under the DSM-5 category of restrictive and repetitive behavior (e.g., rocking, hand flapping, topic perseveration), and these assumptions have been questioned in subsequent research (Bottema-Beutel et al., 2023; Kapp et al., 2019). For example, echolalia—the repetition of the speech of others—has been characterized as problematic and is often targeted for elimination (Blackburn, Tueres, Sandanayake, Roberts, & Sutherland, 2023). Autistic echolalia was initially described as exclusively inwardly focused, produced at random, and a potential barrier to social interaction (Kanner, 1943). However, more recent research has shown that autistic echolalia is an important communicative resource. Careful observational analyses have shown how many echolalic utterances serve dynamic and sophisticated interactional purposes (Pruccoli, Spadoni, Orsenigo, & Parmeggiani, 2021; Sterponi & Shankey, 2014). As such, eliminating them from autistic people's repertoires would likely inhibit their participation in social life, especially for those autistic children who rely exclusively on echolalia for communication.

This example suggests researchers and practitioners should exercise care when determining whether any given behavior is actually a problem and whether interventions to reduce it would cause unintended harm with little or no potential benefit. In the context of intervention research, this care should extend to the measurement tools researchers select to track behaviors they aim to reduce. It is important that such tools offer valid assessments of the subset of behaviors that pose real, immediate, or long-term risks to autistic people and their families or are manifestations of distress, and should exclude any behaviors that are simply manifestations of being autistic.

Theoretical explanations for problem behavior

There are a variety of field-specific explanations as to why problem behavior may occur. According to behaviorist theory, any behavior can be causally linked to its antecedents and consequences (Lovaas & Smith, 1989). Antecedents are environmental

© 2025 The Author(s). *Journal of Child Psychology and Psychiatry* published by John Wiley & Sons Ltd on behalf of Association for Child and Adolescent Mental Health.

14697610, 0, Downloaded from https://acamh.onlinelibrary.wiley.com/doi/10.1111/jcpp.14177 by Turkey Cochrane Evidence Aid, Wiley Online Library on [23/09/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

'stimuli' preceding the behavior that prompt its occurrence, such as a demand from a caregiver. Consequences can make the behavior more or less likely to occur again, such as when a child's behavior is followed by relief/removal from a stressful situation. Some behaviors are considered reinforcing in and of themselves, such as stereotypic behavior (Rapp & Vollmer, 2005). That is, producing the behavior is its own reward, and it is not reliant on external reinforcement.

Other approaches to problem behavior consider a broader and more complex array of phenomena to explain the occurrence of behavior that could be problematic. For example, theoretical frameworks from mental health fields consider behaviors like self-injury to reflect complex emotional, cognitive, relational, and physiological processes that aren't reducible to temporally close environmental stimuli (Maddox, Trubanova, & White, 2017). For young autistic children, difficulties with emotional and sensory regulation are theorized to account for some instances of aggression and self-injury, which may be strategies for coping with emotional and sensory overload and/or overreactivity (Sivathasan et al., 2024). Importantly, this research has shown that many of the most concerning problem behaviors shown by autistic children, such as aggression, also appear in young children who do not have developmental or mental health diagnoses and are not likely explained by autism diagnoses per se, but by emotional reactivity (Sivathasan et al., 2024). Additionally, distress from co-occurring conditions, such as digestive problems or sleep disorders, may cause physiological distress that can manifest in behaviors like self-injury or aggression (Leader, Tuohy, Chen, Mannion, & Gilroy, 2020). Finally, anxiety and depression, which can be more likely to occur in autistic children because of a predisposition for mental health concerns and/or as a response to social contexts in which autistic people are not accommodated and experience stigma, can give rise to coping behaviors that can appear maladaptive or problematic (Han, Scior, Avramides, & Crane, 2022).

The theoretical approach one takes to understanding why problematic behavior can occur is linked to conceptualizations of the behaviors themselves. 'Tantrums' or 'outbursts' are traditionally characterized in behaviorist research as tactics to gain control of the environment, secure others' attention, escape from discomfort, gain access to a desired item or activity, or self-stimulate (Ala'i-Rosales et al., 2019; Tarbox et al., 2009). However, research and advocacy by autistic people have offered a counterframing of these behaviors as meltdowns, shutdowns, and burnout. Autistic advocates describe meltdowns as uncontrollable behavior due to sensory overload, prolonged stress, or an inability to be understood by those around them (Phung, Penner, Pirlot, & Welch, 2021; Schaber, 2014). Conversely, shutdowns involve withdrawal from surroundings and a feeling of detachment or of being 'frozen' because of sensory or emotional overwhelm (Belek, 2019: Phung et al., 2021). Burnout is described as a manifestation of chronic exhaustion following extensive efforts to cope with living in sensory and social environments not designed with autistic people in mind, followed by a loss of skills and reduced tolerance for stimulation (Raymaker et al., 2020). These reconceptualizations invite systemic changes in social and built environments, so there is greater social understanding for autistic people's differences, and the sensory landscape is better aligned with autistic people's sensory and emotional needs.

Intervention to reduce problem behavior

Conceptualizations of problem behavior are also connected to the intervention strategies researchers develop to support autistic children who show such behavior. Several interventions for young autistic children purport to have mechanisms of action for influencing problem behavior, but behaviorally based interventions perhaps provide the most direct description of their influence on behavior change. Given behaviorist logic considers all behavior to bear functional relations with events preceding and following the behavior, strategies for changing behavior considered undesirable involve adjusting relevant antecedents or consequences to reduce the likelihood that the behavior will occur in the future. Interventions to manipulate the consequences of a behavior can involve punishments (Ayvaci et al., 2024) such as removing something desirable when the behavior occurs or introducing an aversive stimulus, and/or they can involve reinforcements such as giving rewards when the child engages in a behavior considered acceptable in place of the problem behavior. Antecedent interventions could involve changing the child's routine to be more predictable, providing choice, and allowing greater access to preferred activities and materials so problem behaviors are no longer relevant to the child (if, e.g., the problem behavior was a strategy to avoid dispreferred tasks or to access desired items).

Other approaches, such as the TEACCH approach, involve arranging the child's environment to cater to their sensory and regulatory needs, including improving the organizational flow of activities to visually clarify activity steps and when it is finished, and reducing environmental stimulation (e.g., putting away toys that aren't being used), which may reduce distraction, task-based overwhelm, and sensory overload (Turner-Brown, Hume, Boyd, & Kainz, 2019). Interventions incorporating developmental principles tend to consider behavior as manifestations of developmental processes and may seek to directly target developmentally proximal outcomes such as play and communication, which could have downstream effects on children's prosocial behavior and emotional regulation, or vice versa

© 2025 The Author(s). *Journal of Child Psychology and Psychiatry* published by John Wiley & Sons Ltd on behalf of Association for Child and Adolescent Mental Health.

14697610, 0, Downloaded from https://acamh.onlinelibrary.wiley.com/doi/10.1111/jcpp.14177 by Turkey Cochrane Evidence Aid, Wiley Online Library on [23/09/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

(Boyd et al., 2018; Harrop, Gulsrud, Shih, Hovsepyan, & Kasari, 2017). Sensory interventions aim to support children in regulating sensory input and modulating arousal to reduce behaviors produced by the child that are purported to be manifestations of sensory dysregulation (Case-Smith, Weaver, & Fristad, 2015), while animal-assisted interventions are designed to facilitate nonjudgmental animal–human social interaction that may be difficult to achieve between autistic children and humans (O'Haire, 2013).

Study aims

In this secondary analysis of Project AIM, a previously conducted systematic review and meta-analysis (Sandbank et al., 2023), we posed questions about the studies, the individual variables, and the measures used for each variable. Previous meta-analyses have shown this group of intervention studies has risks of bias that do not allow for a determination about intervention efficacy for most outcomes; therefore, our study will not address questions regarding whether interventions are efficacious in altering behavior considered problematic. Instead, at the study level, we wanted to determine the proportion of Project AIM studies that included a problem behavior variable, the types of interventions used to address problem behavior, the rationale provided by the authors for targeting problem behaviors, the strategies they used to conceptualize problem behavior (e.g., offering a list of examples, giving a conceptual definition,[1] or describing the impacts of the behavior), and whether or not the study indicated addressing problem behavior was a primary or secondary purpose of the study. We posed the latter question because we assumed some measures of problem behavior may have been included incidentally and were not the main focus of the study or intervention.

At the variable level, we sought to determine the specific measures used to assess problem behavior, the data collection strategies implemented for each variable (e.g., parent report, researcher observation), and the proportion of variables that were assessed with measures that had undergone at least some validation (e.g., an 'off-the-shelf' instrument) and how many were created by the researchers for the purpose of the intervention study (e.g., a researcher-derived coding protocol of observation sessions).

Finally, we also sought to further characterize the measures most used to assess problem behavior across the studies. We wanted to know what psychometric information was available for each instrument and whether the measure tapped behaviors that had a high potential for harm (e.g., aggression, self-injury), behaviors that are nonnormative but do not have potential for immediate harm (e.g., stereotypic behavior such as tapping or vocal play), or a combination thereof.

Methods

Studies were selected and coded from the corpus of studies gathered for the second iteration of Project AIM, a comprehensive systematic review and meta-analysis of nonpharmacological, group-design intervention studies conducted with autistic children up to age 8 (Sandbank et al., 2023). The coding manual, coding spreadsheet, and a complete reference list of included studies are available on the Open Science Framework (OSF; https://osf.io/g57aj/?view_only=2feae004062043e6bdcd7c4cad641daf).

Project AIM search and screening procedures

The search date for the most recent iteration of Project AIM was November 17, 2021, and included searches from the following databases: Academic Search Complete, CINAHL Plus with full text, Education Source, Educational Administration Abstracts, ERIC, Medline, Proquest Dissertations and Theses, PsycINFO, Psychology and Behavioral Sciences Collection, and SocINDEX with full text. We include a table of search terms in Appendices S1 and S2. A total of 6,427 records generated from this search were screened. We sought gray literature (in addition to the unpublished dissertations and theses generated from the previously listed databases) by searching the journal Trials for published protocols and ClinicalTrials.gov with the search term 'autism' to identify potentially relevant registered but unpublished clinical trials, which yielded 168 potentially relevant trials (n = 168).

Thirteen independent screeners double screened abstracts associated with each record using the web application abstrackr. Full texts were examined for any record flagged as potentially eligible by at least one screener. Studies were considered eligible if they were published in English between 1970 and November 2021 (the most recent search date); they were experimental (i.e., a randomized controlled trial) or quasi-experimental group-design studies that included an intervention and a control or comparison group; they reported a simple majority of participants had autism; they included participant samples with an average age <8 years (<96 months). Articles identified through this search were then combined with studies identified in the previous Project AIM search conducted in November of 2017. Authors of any article that did not include the quantitative information necessary for inclusion in the meta-analysis were contacted, and articles were included in the event the necessary information was supplied to the study team. See the original study (Sandbank et al., 2023) for the PRISMA diagram illustrating the flow of records through the search and selection process.

Secondary selection procedures

To extract studies that targeted behavior for reduction, we selected all Project AIM studies that included outcomes categorized as 'challenging behavior' or 'RRBs'. We also selected studies that included outcomes categorized as 'social emotional' if they indexed behaviors targeted for reduction. Outcomes were categorized during the initial process of in-depth coding for the original study, and outcome categorizations were based on examination of both the measure used and the developers' described purpose for the measure (including scales and subscales). We included these latter two categories because, while not all RRB measures describe RRBs as 'problem behaviors', when intervention researchers use them as outcome measures in intervention studies and seek to decrease their occurrence, they are essentially

© 2025 The Author(s). *Journal of Child Psychology and Psychiatry* published by John Wiley & Sons Ltd on behalf of Association for Child and Adolescent Mental Health.

14697610, 0, Downloaded from https://acamh.onlinelibrary.wiley.com/doi/10.1111/jcpp.14177 by Turkey Cochrane Evidence Aid, Wiley Online Library on [23/09/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

construing these behaviors as problematic for autistic children. Additionally, RRBs (especially 'stereotypic behavior') have been subsumed under the problem behavior category in problem behavior literature (e.g., Rapp & Vollmer, 2005). If full-scale measures were described by the developer as assessing a problem behavior construct, we included total scores for the measure in addition to all subscales. If measures were described by the developer as assessing a problem behavior construct in addition to other constructs, we excluded total scores and subscales for which the majority of items did not describe behaviors; for example, if they described internal states (e.g., mood, depression, anxiety) and physiological symptoms (e.g., nausea, headaches). The only exception to this rule was if a study included only total scores, and the authors described the outcome as measuring problem behavior.

Coding of primary studies

The first and second authors independently coded 100% of included studies for the additional variables examined in this secondary analysis. Mean percent agreement between coding was 95% and ranged from 87% to 100% across coding categories. All discrepancies were resolved through consensus between the two coders. Information such as participant mean age, percent male, intervention type, and names of measures was extracted from previous coding. Intervention type categories included behavioral, naturalistic developmental behavioral interventions (NDBIs), developmental, sensory-based intervention, sensory integration therapy, music, TEACCH, animal-assisted therapy, cognitive behavior therapy, computer-based interventions, and 'other' (any intervention that could not be categorized into one of these ten categories).

Purpose. Studies were coded for whether they indicated reducing or eliminating specific behaviors was a primary aim of the study or a direct focus of the intervention, a secondary aim of the study, or did not describe behavior reduction as a focus of the study or described aims too vaguely to categorize (e.g., the study described 'improvement in autistic symptoms').

Rationale. Text describing the authors' rationale for including problem behavior as an outcome was first cut and pasted into a spreadsheet. Once the text for rationale was agreed on by the two coders, it was then coded for whether it referenced negative effects on others, concurrent negative effects on autistic children, long-term negative effects on autistic children, and/or increase in stigma. Note that these categories are not mutually exclusive.

Conceptualization strategy. Similar to rationale coding, we first gathered text excerpts describing the authors' conceptualization of problem behavior from the introduction of primary studies and copied them into a spreadsheet. Note that we did not assess the methods section because we were interested in the framework for conceptualizing, and not operationalizing, problem behavior constructs. Once the text excerpts for conceptualization were agreed on by the two coders, we used procedures described in Bottema-Beutel et al. (2024) and coded this text for whether the authors provided a list of examples, offered a conceptual definition, described the impacts of the behaviors, or some combination of these strategies.

'Off-the-shelf' versus bespoke measures. Each variable was coded for whether it was measured by an instrument that had undergone psychometric assessment beyond intracoder reliability (e.g., convergent validity, factorial validity) or not. We considered all measures to have undergone some validation to be an 'off-the-shelf' measure. Measures created by researchers for the purpose of the intervention study and that had not undergone previous validation beyond intracoder reliability (e.g., coding schemes of parent–child interaction sessions) were considered 'bespoke'.

Data collection strategy. Each variable was coded for data collection procedures. At initial coding, categories included parent report, researcher-scored observation, task-based measure (e.g., participant completion of a worksheet that was not observationally scored), and 'other'. Following initial coding, the 'other' category was subdivided into clinician/trained evaluator observations (who were not reported to be a member of the research team), autistic informants with the help of parents, evaluation of clinical notes, and not specified.

Analysis of most commonly used measures

Following the coding of primary studies, a subset of measures was selected that had been used in five or more primary studies. Ten measures fit this criterion, and all of these were off-the-shelf, standardized instruments. Manuals and instruments were gathered for each of these measures with the help of the university test collection librarian (also the third author).

Psychometric information. Psychometric information was extracted from the instrument manuals.[2] We also conducted basic Google Scholar searches to gather additional psychometric studies that may have been conducted after the publication of each manual. For these searches, the name and abbreviated name of each measure, along with the terms 'validity', 'reliability', and 'psychometric', were entered into the search bar, and the relevant dates were restricted to the publication date of the manual to the present year (2024). To locate studies reporting on instrument validation procedures specifically on autistic populations, the above search was repeated with the addition of the term 'autism'. Titles of the first six pages of results, except for dissertations and theses, were inspected to determine whether any results were relevant. A list of instrument manuals and validation studies is available on OSF. Adapting from Higgins, Kaplan, Deschrijver, and Ross (2024), we summarized psychometric categories separately for autistic and nonautistic populations, including: (a) criterion validity, (b) divergent validity, (c) known group validity, (d) internal consistency, (e) factorial validity, and (f) interrater reliability.

Item-level analysis of harmful versus nonnormative but not harmful behaviors. Measurement items or observation protocols were then coded to determine whether they included items or coding procedures that assessed behavior that was: (a) nonnormative but not harmful, (b) harmful to self, (c) harmful to others, and/or (d) harmful to property. We defined harm to self or others as bodily injury (e.g., cutting or bruising the skin, breaking bones), high risk of bodily injury or trauma (e.g., eloping), or psychological harm (e.g., teasing or threatening). We defined damage to property as breaking, hitting, damaging, stealing, or defacing objects.

Results

Study-level findings

See Table 1 for a complete list of included studies, the purpose of the study, rationale for addressing problem behavior, and strategies for conceptualizing problem behavior. Of the 289 reports included in the

© 2025 The Author(s). *Journal of Child Psychology and Psychiatry* published by John Wiley & Sons Ltd on behalf of Association for Child and Adolescent Mental Health.

14697610, 0, Downloaded from https://acamh.onlinelibrary.wiley.com/doi/10.1111/jcpp.14177 by Turkey Cochrane Evidence Aid, Wiley Online Library on [23/09/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 1 Intervention type, purpose, rationale, and conceptualization strategy by study

Study	Mean age (months)	Mean mental age	% Male	Intervention type	Purpose	Rationale[a]	Conceptualization[b]
Aldred et al., 2004	47.60	NA	89.30	Developmental	Vague	–	–
Argumedes et al., 2021	51.60	NA	78.30	Behavioral	Primary	2	1, 2, 3
Ben-Itzchack & Zachor, 2021	49.20	NA	83.56	Animal-assisted Therapy	Vague	–	1
Bentenuto et al., 2020	41.56	29.62	NA	NDBI	Vague	–	–
Boyd et al., 2014	48.20	NA	83.10	Other, TEACCH	Vague	–	–
Boyd et al., 2018	49.60	NA	86.30	NDBI	Secondary	1, 3	–
Brian et al., 2017	25.20	NA	75.80	NDBI	Vague	–	–
Cai et al., 2020	57.54	NA	86.67	Other	Vague	–	–
Carruthers et al., 2021	45.00	16.20	90.80	Developmental	Vague	–	–
Charman et al., 2021	80.04	NA	80.60	Behavioral	Primary	–	–
Clionsky, 2012	61.40	NA	80.00	Other, Other	Primary	–	–
Cook et al., 2019	65.40	NA	87.10	Cognitive Behavior Therapy	Vague	1, 3	–
Dawson et al., 2010	23.50	NA	77.10	NDBI	Vague	–	–
De Korte et al., 2020	75.94	NA	84.09	Computer-based	Vague	–	1, 3
D'Elia et al., 2014	49.20	NA	80.00	TEACCH	Secondary	–	–
Devescovi et al., 2021	29.58	NA	84.21	NDBI	Vague	–	–
Dimachkie, 2020	31.48	NA	82.70	NDBI	Primary	1, 3	2, 3
Divan et al., 2019	64.98	NA	87.50	Developmental	Primary	1, 2	
Drew et al., 2002	22.50	NA	79.20	NDBI	Primary	–	–
Duifhuis et al., 2017	69.50	NA	83.30	NDBI	Secondary	–	–
Durand et al., 2012	51.10	NA	85.70	Other	Primary	1, 2, 3	1, 3
Eikeseth et al., 2002	65.70	NA	76.00	Behavioral	Vague	–	–
Eikeseth et al., 2007	65.70	NA	76.00	Behavioral	Primary	–	1
Estes et al., 2015	23.50	NA	76.90	NDBI	Primary	–	–
Fava et al., 2011	48.20	NA	86.40	Behavioral	Primary	–	–
Fletcher-Watson et al., 2016	49.60	NA	79.60	Computer-based	Vague	–	–
Furukawa et al., 2018	62.90	NA	81.38	Other	Primary	1, 2	1, 3
Gengoux et al., 2021	60.40	NA	95.40	NDBI	Vague	–	–
Gepner et al., 2021	68.20	NA	86.90	Computer-based	Primary	–	–
Grahame et al., 2015	61.50	NA	86.70	Behavioral	Primary	1, 2, 4	1, 2, 3
Green et al., 2010	45.00	26.20	90.80	Developmental	Vague	1, 2, 3	–
Green et al., 2022	48.00	NA	79.40	Developmental	Secondary	–	–
Hemdi & Daley, 2017	61.00	NA	NA	Other	Primary	1	–
Holzinger et al., 2019	43.31	NA	100.00	NDBI	Vague	–	–
Howlin & Rutter, 1989	75.40	NA	100.00	NDBI	Primary	–	1
Iadarola, Levato et al., 2018	57.00	NA	79.00	Behavioral	Primary	1	1, 3
Iadarola, Shih et al., 2018	85.10	NA	87.30	Behavioral	Secondary	3	1, 3
Ichikawa et al., 2013	64.00	NA	81.80	TEACCH	Primary	1	
Jónsdóttir et al., 2018	37.50	NA	86.66	Behavioral	Vague	–	–
Jouen et al., 2017	83.80	NA	100.00	Computer-based	Vague	3	–
Kitzerow et al., 2020	63.00	49.80	95.00	NDBI	Vague	–	–
Kovshoff et al., 2011	37.00	22.80	NA	Behavioral	Vague	–	–
Kubo et al., 2021	89.20	NA	76.70	Other	Primary	1, 2	–
LaGasse, 2014	91.00	NA	76.50	Music	Vague	–	–
Lau, 2017	46.20	NA	88.89	Other	Primary	1	1, 3
Leadbitter et al., 2018				Developmental	Vague	–	–
Leaf et al., 2017	56.40	NA	NA	Behavioral	Primary	–	–
Lindgren et al., 2020	52.20	NA	84.20	Behavioral	Primary	2	1, 3

(continues)

© 2025 The Author(s). *Journal of Child Psychology and Psychiatry* published by John Wiley & Sons Ltd on behalf of Association for Child and Adolescent Mental Health.

14697610, 0, Downloaded from https://acamh.onlinelibrary.wiley.com/doi/10.1111/jcpp.14177 by Turkey Cochrane Evidence Aid, Wiley Online Library on [23/09/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 1 (continued)

Study	Mean age (months)	Mean mental age	% Male	Intervention type	Purpose	Rationale[a]	Conceptualization[b]
Lopez et al., 2019	63.80	NA	88.50	Other	Vague	–	–
Magaña et al., 2020	63.70	NA	87.00	Other	Secondary	–	–
Mahoney & Solomon, 2016	49.80	NA	84.00	Developmental	Primary	–	–
Mazurek et al., 2020	49.40	NA	80.10	Behavioral, NDBI, NDBI, NDBI	Primary	–	–
McCalla, 2014	57.30	NA	89.70	Behavioral	Vague	–	1, 2
McEachin et al., 1993	33.50	NA	NA	Behavioral	Primary	–	1
McGregor et al., 2020	50.10	NA	71.30	Other	Primary	2, 3	1, 2, 3
Murdock et al., 2014	51.50	NA	86.70	Sensory-based	Primary	–	–
Page, 2012	90.50	NA	53.80	Animal-assisted Therapy	Primary	–	–
Pan, 2010	86.80	NA	100.00	Other	Vague	–	–
Peters-Scheffer et al., 2013	62.50	23.40	90.00	Behavioral	Primary	–	1
Pickles et al., 2016	45.00	16.20	90.80	Developmental	Primary	1, 3	
Pravej et al., 2009	55.90	NA	81.70	Sensory-based	Primary	–	–
Qiu et al., 2021	52.20	NA	75.00	Other	Primary	–	–
Rabeyron et al., 2020	61.60	NA	86.10	Music	Secondary	–	–
Reitzel et al., 2013	60.20	NA	NA	Behavioral	Primary	–	1
Remington et al., 2007	37.00	22.80	NA	Behavioral	Primary	–	–
Roberts et al., 2011	42.60	NA	90.50	NDBI	Vague	–	–
Rogers et al., 2012	21.00	NA	77.50	NDBI	Vague	–	–
Rogers et al., 2021	23.44	14.52	77.27	NDBI, Behavioral	Vague	–	–
Saint-Georges et al., 2020	84.71	NA	84.72	Other	Primary	–	–
Sallows & Graupner, 2005	33.70	NA	82.60	Behavioral	Vague	–	–
Sampanthavivat et al., 2012	70.60	NA	93.10	Other	Vague	–	–
Scarpa & Reyes, 2011	67.70	NA	81.80	Cognitive Behavior Therapy	Vague	2	–
Schaaf et al., 2014	71.80	NA	81.25	SIT	Secondary	1, 2	2, 3
Scheithauer et al., 2021	80.80	NA	86.90	Behavioral	Primary	2, 3	1, 2, 3
Schottelkorb et al., 2020	NA	NA	82.60	Other	Primary	2, 3	1, 3
Schwartzman et al., 2022	79.80	NA	82.30	Other	Secondary	–	–
Scudder et al., 2019	67.40	NA	89.50	Other	Primary	2, 3	1, 3
Sharp et al., 2019	58.70	NA	84.20	Other	Primary	1, 2, 3	2, 3
Smith et al., 1997	37.00	NA	90.50	Behavioral	Vague	–	–
Smith et al., 2000	35.90	21.70	82.10	Behavioral	Vague	–	–
Srinivasan et al., 2015	89.28	NA	91.60	Computer-based	Primary	1, 2, 3, 4	1, 3
Suzuki et al., 2014	54.00	NA	NA	Other	Secondary	1, 2, 3, 4	3
Tanet et al., 2020	84.71	NA	84.72	Other	Secondary	–	–
Tellegen & Sanders, 2014	68.10	NA	85.90	Other	Primary	–	–
Thompson et al., 2013	45.40	NA	82.60	Music	Secondary	–	–
Tonge et al., 2014	48.10	NA	82.90	Behavioral, Other	Primary	–	–
Valeri et al., 2019	51.60	NA	79.40	NDBI	Secondary	–	–
Vivanti et al., 2014	41.20	20.70	87.70	NDBI	Vague	–	–
Vivanti et al., 2019	25.55	NA	61.36	NDBI	Secondary	–	–
Wan et al., 2021	90.90	NA	85.00	Other	Primary	2	1, 3
Wang et al., 2020	59.10	NA	84.80	Other	Vague	2	3
Whitehouse et al., 2017	39.80	NA	78.80	Computer-based	Secondary	–	–
Whittinham et al., 2009	70.90	NA	79.70	Behavioral	Primary	–	–

(continues)

© 2025 The Author(s). *Journal of Child Psychology and Psychiatry* published by John Wiley & Sons Ltd on behalf of Association for Child and Adolescent Mental Health.

14697610, 0, Downloaded from https://acamh.onlinelibrary.wiley.com/doi/10.1111/jcpp.14177 by Turkey Cochrane Evidence Aid, Wiley Online Library on [23/09/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 1 (continued)

Study	Mean age (months)	Mean mental age	% Male	Intervention type	Purpose	Rationale[a]	Conceptualization[b]
Williams et al, 2020	67.90	NA	70.70	Other	Primary	1	1
Wong & Kwan, 2010	26.50	17.90	94.10	Other	Vague	1	–
Young & Posselt, 2012	NA	NA	NA	Computer-based	Vague	–	–
Yu & Zhu, 2018	53.43	NA	0.79	NDBI	Vague	–	–
Zachor et al., 2016	63.50	NA	78.40	Other	Vague	–	–
Zanobini & Silvano, 2019	66.70	NA	76.00	Other	Primary	–	–
Zeng et al., 2021	54.55	NA	78.33	TEACCH	Primary	–	–
Zeng et al., 2022	63.70	NA	88.10	Other	Primary	–	–
Zhou et al., 2018	73.40	NA	70.70	NDBI	Primary	–	–

NDBI, Naturalistic, Developmental, Behavioral Interventions; SIT, Sensory Integration Therapy.
[a]1 = List of Examples, 2 = Definition, 3 = Description of Impacts.
[b]1 = Effects on Others, 2 = Concurrent Child Effects, 3 = Long-term Child Effects, 4 = Increases in Stigma.

most recent iteration of Project AIM, 102 (35%) included at least one outcome focused on the reduction of behavior(s). We also note that most studies (>80%) were published in the 10 years prior to the most recent search date. These percentages are consistent with the proportion of all studies in our dataset that were published during that time frame, suggesting that there has been no decrease over time in group-design autism intervention researchers' attempts to target behaviors for reduction.

Intervention type. All intervention categories in Project AIM were represented among the studies that measured a behavior or set of behaviors targeted for reduction, and all except cognitive behavior therapy had at least one study indicating behavior reduction was a primary or secondary purpose of the study or intervention. 'Other' interventions most frequently included problem behavior outcomes, followed by behavioral interventions and NDBI interventions. Most intervention types had approximately equal distributions of studies listing problem behavior as a primary/secondary focus of the study or intervention versus studies only vaguely describing problem behavior as a purpose of the study, except for 'other' and behavioral interventions. These two intervention types had more than twice as many studies describing the reduction of problem behavior as a primary or secondary purpose of the study or intervention as compared to studies only vaguely referencing this purpose. See Table 2 for tallies of studies within intervention type and primary/secondary versus vaguely stated purpose.

Purpose. Of these, 48 studies (47%) reported addressing problem behavior as a primary purpose of the intervention and/or trial, and an additional 15 studies (15%) reported addressing problem behavior as a secondary purpose. The remaining 39 studies (38%) either did not indicate that reducing problem behavior was a purpose of the study/intervention or described the study purposes too vaguely to categorize.

Table 2 Number of studies by intervention type and purpose

Intervention	# of Primary/secondary purpose studies	# of Vague purpose studies
Other	23	8
Behavioral	17	8
NDBI	12	12
Developmental	4	4
TEACCH	3	1
Computer-based Intervention	3	4
Sensory-based Intervention	2	0
Music	2	1
Sensory Integration Therapy	1	0
Animal-assisted Therapy	1	1
Cognitive Behavior Therapy	0	2
Total	68	41

NDBI, Naturalistic Developmental Behavioral Interventions.

Rationale. Thirty-three studies (33%) provided a rationale for addressing problem behavior. Of these, 21 (64%) reported rationales describing the effects of problem behavior on others (e.g., parents, teachers, family members, society), 20 (61%) reported rationales describing the immediate effects of problem behavior on autistic children (e.g., decreased social participation), 13 (39%) reported rationales describing long-term consequences of problem behavior on autistic children (e.g., poor mental health, decreased employment opportunities, death), and 3 (9%) reported rationales referencing increased stigma. Note that most studies providing any rationale provided more than one; thus, the total number

© 2025 The Author(s). *Journal of Child Psychology and Psychiatry* published by John Wiley & Sons Ltd on behalf of Association for Child and Adolescent Mental Health.

14697610, 0, Downloaded from https://acamh.onlinelibrary.wiley.com/doi/10.1111/jcpp.14177 by Turkey Cochrane Evidence Aid, Wiley Online Library on [23/09/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

across these rationale categories is greater than 33. Twenty-seven (82%) of the studies that provided a rationale for addressing problem behavior listed addressing the behaviors as a primary or secondary purpose of the intervention.

Conceptualization strategy. Twenty-nine studies (28%) offered a conceptualization of problem behavior. Of these 29 studies, 8 (28%) provided a list of examples of problem behavior; 2 (7%) described the impacts of the behavior, 11 (38%) provided a list of examples and described the impacts of problem behavior; 3 (10%) described the impacts of the behavior and provided a conceptual definition of problem behavior; 1 (3%) provided a list of examples and provided a conceptual definition of problem behavior; and 4 (14%) provided a list of examples, offered a conceptual definition, and described the impacts of problem behavior. Across all studies that conceptualized problem behavior using any of the three conceptualization strategies or a combination thereof, only 8 (28%) defined problem behavior, which corresponds to 8% of the total set of studies that included a problem behavior outcome.

Variable-level findings

Across the 102 included studies, there were 437 problem behavior variables.

'Off-the-Shelf' vs. bespoke. Across these variables, there were 36 unique standardized and previously (or concurrently, in the case of one instrument; Leadbitter et al., 2018) validated assessments, 12 bespoke assessment protocols that had not been subject to psychometric assessment beyond inter-coder reliability (e.g., applying researcher-derived coding schemes to participant observations), and two combination protocols that involved bespoke adaptations to, or combinations of, previously validated measures. See Table 3 for the list of validated assessments and subscales that met our criteria for assessing problem behaviors.

Data collection strategy. Assessment methods included 259 (59%) parent reports, 61 (14%) researcher observations, 45 (10%), 44 clinician/trained evaluator observations (10%), 18 (4%) trained nonobservational task raters, 5 (1%), 2 (<1%) autistic informants and their parents, and 2 (<1%) evaluations of clinical notes.

Analysis of Most commonly used measures

Psychometric information. Findings regarding psychometric assessments are presented in Table 4. Two out of ten instruments reported adequate psychometrics in autistic populations across all six categories: the *Autistic Diagnostic Observation Schedule-2* (Lord & Rutter, 2012) and the *Social Responsiveness Scale-2* (Constantino & Gruber, 2012). For the remaining instruments, we located reports indicating that psychometrics were adequate in autistic populations in two to five categories.

Item-level analysis of harmful vs. nonnormative but not harmful. When the 10 measures were subdivided into full scales and subscales, this resulted in a total of 30 sets of items that were assessed for whether the items described nonnormative but not harmful behavior, or described behavior that was harmful to self, others, or property. Of these 30 full scales/subscales, only a single subscale (the *Self-Injurious Behavior* subscale of the *Repetitive Behavior Scale-R*; Lam & Aman, 2007) included only harmful behavior and did not contain any items describing behavior that was nonnormative but not harmful. Sixteen scales/subscales only included items that were nonnormative but not harmful, and the remaining 13 scales/subscales contained items that were a mix of harmful and nonnormative but not harmful items. These determinations by full scale and subscale for each measure, along with relevant examples, are included in Appendix S1.

The protocol for one instrument, the *Clinical Global Impression Scale*, precluded us from conducting our item-level coding procedures, because the instrument requires the caregiver to describe issues about the child that are problematic, and the severity of these issues is then rated by the researcher or clinician. Therefore, the nature of specific problems will vary from participant to participant. This instrument does not specify problematic to mean harmful, so this scale could also result in a mix of harmful and nonnormative but not harmful behaviors.

Discussion

In this secondary systematic review, we examined study-level, variable-level, and measure-level aspects of studies aiming to reduce behavior(s) exhibited by young autistic children. We found that most studies that met our inclusion criteria indicated reducing behaviors was a primary or secondary purpose of the study/intervention, and that nearly every intervention type had at least one study expressing this purpose. Of interventions classifiable by type, behavioral and NDBI interventions were represented in the largest number of studies that included a problem behavior outcome and that described reducing such behavior as a primary or secondary purpose of the intervention. This suggests to us that a range of theoretical orientations are aligned with targeting problem behavior, but the strongest alignment may be with behaviorally oriented approaches.

© 2025 The Author(s). *Journal of Child Psychology and Psychiatry* published by John Wiley & Sons Ltd on behalf of Association for Child and Adolescent Mental Health.

14697610, 0, Downloaded from https://acamh.onlinelibrary.wiley.com/doi/10.1111/jcpp.14177 by Turkey Cochrane Evidence Aid, Wiley Online Library on [23/09/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 3 Number of studies and included subscales for validated measures

Name of instrument	# of Studies	Full measure or subscales only?
Aberrant Behavior Checklist	11	Full Scale
ADHD Rating Scale	1	Full Scale
Assessment of Concerning Behavior	1	Full Scale
Autism Behavior Checklist	1	Body and Object Use
Autism Diagnostic Interview- Revised	4	Restrictive and Repetitive Behavior
Autism Diagnostic Observation Schedule-2	18	Restrictive and Repetitive Behavior
Autism Spectrum Rating Scales	1	Unusual Behavior, Behavior Regulation
Behavior Problems Inventory	1	Full Scale
Behavior Rating Inventory of Executive Functioning	1	Behavior Regulation
Behavioral Flexibility Rating Scale-Revised	2	Full Scale
Brief Autism Mealtime Behavior Inventory	1	Full Scale
Brief Observation of Social Communication	3	Restrictive and Repetitive Behavior
Child Behavior Checklist	16	Attention Problems, Rule-breaking behavior, Aggressive behavior, Internalizing, Externalizing
Childhood Autism Rating Scale	2	Bizarre Use of Body Movement and Persistence of Stereotypes, Peculiarities in Relating to Nonhuman Objects, Resistance to Environmental Change, Activity Level
Clinical Global Impressions Scale	5	Full Scale
Connors Teachers Rating Scale	1	Full Scale
Developmental Behavior Checklist	7	Full Scale
Eyberg Child Behavior Inventory	5	Full Scale
Home Situations Questionnaire- Autism Spectrum Disorder	1	Full Scale
Improvement in Parent-Defined Target Systems - CGI adaptation	1	Full Scale
Observation Schedule for Children with Autism-Anxiety, Behavior and Parenting	1	Child Behaviors the Challenge, Child Compliance
Parenting Stress Index	2	Difficult Child
PDD Behavior Inventory	3	Sensory/Perceptual Approach Behaviors; Aggressiveness; Specific Fears; Arousal Problems
Psychoeducational Profile	6	Problem Behaviors
Repetitive Behavior Scale	8	Full Scale
Repetitive Behaviors Questionnaire	3	Full Scale
Scales of Independent Behavior	4	Problem Behavior
School Function Assessment	1	Behavior Regulation, Compliance with Adult Directives, Task Behavior/ Completion
School Situations Questionnaire	1	Full Scale
School Social Behavior Scales-2	1	Self-Management/Compliance, Hostile/Irritable, Anti-social/ Aggressive, Defiant/Disruptive
Social Communication Questionnaire	2	Restrictive and Repetitive Behavior
Social Responsiveness Scale	13	Characteristic Autistic Preoccupations/Traits
Strengths and Difficulties Questionnaire	4	Conduct Problems, Hyperactivity/Inattention
Target Behavior Vignette	2	Full Scale
Vineland Adaptive Behavior Scales-2	5	Maladaptive Behavior

Only a third of the studies described a rationale for targeting the behaviors the intervention was hypothesized to reduce. These rationales most often included descriptions of purported negative effects of these behaviors on autistic children or their family members. One possible explanation for the low number of studies providing a rationale is that many intervention researchers may select outcome measures because they index behavior associated with autism and take it as a given that the general purpose of interventions is to reduce these types of behaviors without the need for any explicit rationale. A similarly small minority of studies conceptualized problem behavior in such a way that would allow readers to understand the extent to which any given behavior would meet definitional criteria for the behavior types that were targeted for reduction. The murkiness of the problem behavior construct and the specific behaviors it includes mean it is unclear which behaviors produced by autistic children are believed to negatively impact them and why, leaving the veracity of the provided rationales open to question. That is, there may be empirical evidence that problem behavior like aggression has negative impacts on autistic children, and this evidence could be inappropriately used to justify the targeted reduction of a range of unrelated behaviors that researchers subsume under an overall problem behavior construct. For example, Rapp and Vollmer (2005) have suggested 'stereotypy' should be targeted for reduction both because the behavior is assumed to be meaningless and because 'it absorbs

© 2025 The Author(s). *Journal of Child Psychology and Psychiatry* published by John Wiley & Sons Ltd on behalf of Association for Child and Adolescent Mental Health.

14697610, 0, Downloaded from https://acamh.onlinelibrary.wiley.com/doi/10.1111/jcpp.14177 by Turkey Cochrane Evidence Aid, Wiley Online Library on [23/09/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 4 Validity information for commonly used measures

	Criterion validity	Divergent validity	Known group validity	Internal consistency	Factor structure	Interrater reliability
ABC	A, O	A[a], O	–	A	A, O	A, O
ADOS-2	A	A	A, O	A	A	A
CBCL	A, O	–	A, O	A, O	A[b], O	O
CGI[c]	A, O	O	–	O	–	A, O
DBC-2	O, A	O	A[d]	O, A	O	O, A
ECBI	A, O	A, O	O	A, O	–	O
PEP-3	A	A	A, O	A, O	O	A, O
RBS-R	A, O	A	A	A, O	–	A
SRS-2	A, O	A, O	A, O	A, O	A, O	A, O
VABS-III	A,O	A,O	AO	A,O	–[e]	A,O

'A' indicates the measure was validated with autistic participants, 'O' indicates the measure was validated with other, nonautistic participants. ABC, Aberrant Behavior Checklist; ADOS-2, Autism Diagnostic Observation Schedule-2; CBCL, Child Behavior Checklist; CGI, Clinical Global Impressions Scale; DBC, Developmental Behavior Checklist; ECBI, Eyberg Child Behavior Inventory; PEP-3, Psychoeducation Profile-3; RBS, Repetitive Behavior Scale; SRS-2, Social Responsiveness Scale-2; VABS-III, Vineland Adaptive Behavior Scale-III.

[a]The ABC manual reports that there were no associations with IQ for autistic populations, with the exception of the Stereotypic Behavior Subscale.

[b]Medeiros, Mazurek, and Kanne (2017) found that the factor structure of the CBCL was appropriate for children ages 1.5–5, but not for older children.

[c]Although the CGI is frequently used, in part because of its versatility, Beneke and Rasmus (1992) problematize the vagueness of the instrument, 'since "global impression" does not give any more information than the label "food" on atin: i.e., nobody knows the actual content' (p. 171).

[d]There is an adaptation of the DBC, the *DBC Autism Screening Algorithm*, but this instrument was not used in the included studies. There is one study that included 30 autistic children in a total sample of 531 children, but results were not reported separately for the autistic subsample.

[e]The Vineland Manual reports that the three-domain/factor structure is conceptually rooted. Wilkinson, Farmer, Kleiman, and Bal (2023) report they were the first to assess this factor structure in autistic participants and found the three-factor structure to be a poor fit.

too much of an individual's time' (p. 528). However, contemporary research has refuted these notions and suggests that stereotypic behaviors (and other RRBs) can be expressions of joy or a means to cope with an overstimulating environment (Kapp et al., 2019).

Adequate conceptualization of outcome constructs is also important for determining the extent to which a given construct is sufficiently matched to the measurement tools chosen to assess it (Wigham & McConachie, 2014). Because conceptual definitions were only provided in 8% of studies, this kind of assessment is generally not possible in this body of research. However, given that there are large differences in how researchers define problem behavior, or even subtypes of this construct such as stereotypy (see Rapp & Vollmer, 2005) and aggression (see Sivathasan et al., 2024), adequate conceptualizations that include basic definitions are critically important for intervention researchers to communicate to readers. This would allow end users of interventions (e.g., parents, practitioners) to make determinations about whether and how interventions have the potential to impact the kinds of behaviors that are important for autistic children and families.

In the studies we examined, off-the-shelf measures were far more common than bespoke measurement systems created by researchers, which contrast with a previous analysis of problem behavior outcomes in intervention research with transition-age autistic youth in which bespoke measurement systems were most common (Bottema-Beutel et al., 2024). There are several advantages of off-the-shelf measurement tools, including that they have undergone at least some psychometric evaluation, and that they tend to measure generalized constructs that are not overly aligned with intervention targets (see Sandbank, Chow, Bottema-Beutel, & Woynaroski, 2021 for a discussion). Parent reports are currently the most common method of gathering information on autistic children's problem behavior. However, this strategy is problematic for intervention research because caregivers are rarely masked to group assignment, and parent reports are therefore subject to detection bias. That is, intervention effects can be inflated due to the assessor knowing the participant is receiving the intervention (Higgins et al., 2021).

The measurement-level analysis showed the most common instruments reported adequate assessment for at least some aspects of instrument psychometrics, but only 20% of instruments appeared to have been evaluated across all six psychometric categories. However, in our item-level analysis, we found most of these measures intermingle behaviors that are nonnormative but not harmful (e.g., shyness, preoccupation with objects) with behaviors associated with high potential for harm (e.g., setting fires, biting, deliberate self-harm), with only a single exception across 30 measures/subscales. This could be an issue of face validity; although some of these

© 2025 The Author(s). *Journal of Child Psychology and Psychiatry* published by John Wiley & Sons Ltd on behalf of Association for Child and Adolescent Mental Health.

14697610, 0, Downloaded from https://acamh.onlinelibrary.wiley.com/doi/10.1111/jcpp.14177 by Turkey Cochrane Evidence Aid, Wiley Online Library on [23/09/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

scales show adequate psychometric properties across a range of dimensions, it is not clear they are measuring a useful and well-defined superordinate construct. If a study were to claim the intervention showed the desired reduction in scores from these measures, in most cases, it would be impossible to determine whether the intervention is simply encouraging autistic children to mask their autistic traits or whether the intervention is reducing behaviors with clear risks of immediate or long-term harm. It is important for researchers to consider the potential negative effects of interventions that promote masking, as this phenomenon has been associated with longer-term negative outcomes for autistic people (Cook, Hull, Crane, & Mandy, 2021; Kapp, 2014).

At present, many autism researchers are adopting neurodiversity-aligned approaches to intervention, where normalization is no longer viewed as a defensible goal (Leadbitter, Buckle, Ellis, & Dekker, 2021). A recent survey of autistic people and autism community members (e.g., caregivers of autistic children, service providers) indicated that reducing autistic behavior (including nonharmful behaviors that fall under the RRB umbrella) is generally *not* a desired component of early intervention (Bent et al., 2024). More acceptable goals include focusing on mitigating behaviors that could be harmful by focusing on why they occur (e.g., physical or emotional distress) and monitoring other behaviors that could be expressions of distress (e.g., stimming to cope with sensory overload). To meet these goals, however, measures will need to be developed and validated that make these distinctions, and informants will need to be able to determine which behaviors are manifestations of distress, which are neutral, and which are expressions of positive emotions such as happiness and joy (Kapp et al., 2019).

Limitations

There are some limitations that should be considered when interpreting the results of this study. First, as this is a secondary systematic review of a previously conducted meta-analysis, we only examined studies that were included in the original review. These studies were limited to those that were: (a) controlled group experimental designs and (b) focused on interventions for autistic children up to age eight. It is possible study-, variable-, and measurement-level findings will be different for studies examining behaviors targeted for reduction in other disability groups, age ranges, and design types. Notably, most behavior analytic intervention research uses single-case designs, which tend not to use 'off-the-shelf' outcome measures because they typically are not suited for use as high-frequency repeated measures. As such, at least some of our findings are not relevant to that group of studies.

Second, the vagueness of the problem behavior concept means selecting studies relevant to this concept was not clear cut. Some studies used measures for which it was clear that behaviors were being targeted for reduction, but the studies themselves did not describe the reduction of problem behaviors as a purpose of the study. Other studies described the reduction of problem behaviors as a purpose of the study but did not use measures that are marketed as problem behavior measures. We used selection criteria that were designed to capture both types of studies, but more conservative selection criteria would likely yield different results.

Third, our psychometric study did not consist of formal systematic review procedures; we relied on information published in developer manuals and a semisystematic review of Google Scholar results. We also provide validity information published between the dates of the most recent manual publication (since previously published versions are generally not available) through August 2024. However, many studies, and in some cases manuals, were published after the publication of the primary studies in our corpus from Project AIM. Given this discrepancy, our validity findings are most helpful as a guide for future instrument selection and not necessarily as a record of validity information that the original study authors would have had when selecting instruments to measure intervention outcomes. We also did not examine validation practices for the studies that used instruments that were not represented in five or more studies.

Finally, we recognize that our categorization of behaviors as either inherently harmful or not can serve to obscure that some behaviors that are not immediately harmful may become harmful over longer periods of time, either by escalating or by repeatedly disrupting family functioning. Judgments about the potential harmfulness of a given behavior are complex and depend on several factors that caregivers and early childhood intervention clinicians often take into account, and these factors may not readily be apparent to outside observers. We assert that greater consideration of the nuances that distinguish behaviors as reasonable targets for reduction or further monitoring is necessary. However, consensus building across the field and the autistic community on this issue cannot happen unless researchers do more to offer conceptual definitions of specific behaviors, create measures that require consideration of the social contexts in which behaviors do or do not have implications for immediate or later harm, and better articulate the rationale for targeting them for reduction and monitoring.

Conclusions

In light of our findings, we conclude that the construct of 'problem behavior' (and its many

© 2025 The Author(s). *Journal of Child Psychology and Psychiatry* published by John Wiley & Sons Ltd on behalf of Association for Child and Adolescent Mental Health.

14697610, 0, Downloaded from https://acamh.onlinelibrary.wiley.com/doi/10.1111/jcpp.14177 by Turkey Cochrane Evidence Aid, Wiley Online Library on [23/09/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

terminological variations such as aberrant, challenging, or maladaptive behavior, and subvariants such as stereotypy or restrictive and repetitive behavior) is inadequately conceptualized in the context of intervention research for young autistic children. Exacerbating the vagueness of the general construct is the fact that intervention researchers rarely provide a conceptual definition of the specific behavior constructs they plan to target and provide a rationale for targeting behaviors for reduction only some of the time. Likewise, the measures that have been used to assess intervention effects in reducing behaviors are of unclear utility, given most of them mix the assessment of behaviors that have a high likelihood of causing harm with behaviors that are nonnormative but not inherently harmful.

We have several recommendations for researchers in light of our findings. First, interventions should be developed that have clear theories of change that connect the components of the intervention to the measurement outcomes (e.g., a 'program theory'; Skivington et al., 2021). In the case of interventions that aim to reduce specific behaviors, autism researchers need to develop and clearly communicate how and why the intervention will yield change in this domain—noting there have been many missteps in previous intervention research when behavior constructs were not fully conceptualized and/or understood prior to intervention development. Outcome measures should be selected that adequately match appropriate conceptualizations of the behavior domains being targeted. Researchers should avoid selecting measures solely because they are commonly used in autism intervention research and/or are readily available, and should also avoid 'incidental' inclusion of outcome measures that have no connections to theory of change that underlies the intervention.

Our findings also suggest new measures will need to be created in consultation with researchers who have psychometric and content-area expertise (within and beyond autism research), and with autism/autistic communities to determine the utility of targeting any given set of behaviors for reduction via intervention (see Leadbitter et al., 2018 for an example of this approach), and to establish face and content validity of any test items that comprise such a measure (Hardesty & Bearden, 2004). To support this work, foundational research will need to be conducted to better understand, describe, and identify these behaviors, and to appropriately distinguish them from behaviors expressing happiness, curiosity, or other human experiences. These distinctions are important to avoid promoting camouflaging as an intervention outcome. When designing these measures, researchers should consider the social and relational contexts in which behaviors are/are not considered problematic (Bottema-Beutel, 2024; Simplican, 2019). A useful starting point for replacing the 'problem behavior' construct could be 'behaviors that cause harm'. Once new measures are drafted, they should be subject to appropriate psychometric validation with autistic participants before they are used in intervention research.

Some behaviors that are not inherently harmful could also be useful to track as intervention outcomes, such as behaviors that are expressions of distress, behaviors that are associated with decreased community access, or behaviors that caregivers believe will lead to eventual harm even if they are not immediately harmful. However, these potentially harmful behaviors should be clearly distinguished from behaviors that are inherently harmful in the administration and scoring protocols, and intervention researchers should clearly articulate their rationale for tracking these behaviors. Lumping inherently harmful and potentially harmful behaviors together in a single measure, as is commonly done in current practice, conceals important nuances about behavior of which families, autistic people, and researchers need to be aware.

Measurement protocols should also be developed that can be administered such that assessors are masked to treatment status. This could mean that improved caregiver report measures could be developed that are administered in the context of active-treatment controlled trials, in which caregivers in all groups are masked as to which condition is hypothesized to be superior. In instances where active-treatment controls are not preferable, other measures of behaviors researchers aim to decrease could be developed that involve observations of autistic children in meaningful contexts (e.g., home or school observation protocols) that can be administered and scored by assessors masked to group assignment. Alternatively, researchers might elect to use caregiver report measures, but ask participating families to select a caregiver who regularly interacts with the child (i.e., a preschool teacher, an extended family member) to remain masked to intervention assignment throughout the course of the study.

Finally, our findings have implications beyond autism intervention research, as the measures we examined are used in intervention research with numerous other populations, such as children with attention deficit hyperactivity disorder (Aldabbagh, Glazebrook, Sayal, & Daley, 2024) and children with intellectual disabilities (Hassiotis & Rudra, 2022). There have been some efforts at rethinking conceptualization of these behaviors in research with populations with intellectual disabilities, for example, by referring to them as 'behaviors that challenge' (primarily in the United Kingdom), to clearly acknowledge that the perspectives of people other than the person exhibiting the behavior are being prioritized, and to convey that the 'challenge' is a result of the interaction between the behavior and the environment and is not located within the person (National Collaborating Center for Mental

© 2025 The Author(s). *Journal of Child Psychology and Psychiatry* published by John Wiley & Sons Ltd on behalf of Association for Child and Adolescent Mental Health.

14697610, 0, Downloaded from https://acamh.onlinelibrary.wiley.com/doi/10.1111/jcpp.14177 by Turkey Cochrane Evidence Aid, Wiley Online Library on [23/09/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Health, 2015). However, even this terminology has been critiqued for insufficiently directing attention to the social conditions in which such behaviors occur and are conceptualized as challenging, and changes in terminology have not been accompanied by calls to substantively change how these behaviors are assessed. Intervention research in these fields may also need to reconsider how this construct is conceptualized and measured.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article:

Appendix S1. PICOS (Population, Intervention, Comparison, Outcomes, and Study) search strategy.

Appendix S2. Item-level analysis of harmful vs. not harmful items for included full scales and subscales.

Acknowledgments

This project was partially funded by NIH grant # 1R01DC020186 and the Argyelan Family Fund, administered by the Lynch School of Education at Boston College.

K.B.-B. has previously received fees for consulting with school districts on intervention practices for autistic children and teaches courses on autism interventions in her role as a professor of special education. She has also accepted speaker fees to discuss her work on research quality, adverse events, and researcher conflicts of interest as they pertain to autism intervention research. She also receives royalties for a co-edited book titled *Clinical Guide to Early Interventions for Children with Autism*, published by Springer. Y.S. previously completed a Master's program in Applied Behavior Analysis and has received compensation for work as a behavioral interventionist in several inclusion classrooms, working with autistic individuals from early childhood through adulthood across different school settings. S.L. was formerly affiliated with an entity that trained students to become board-certified behavior analysts. She was previously employed by the TEACCH Autism Program and served as an interventionist on an intervention developed at TEACCH for autistic transition-age youth. T.W. is the parent of an autistic child; has previously been paid to provide traditional behavioral, naturalistic developmental behavioral, and developmental interventions to young children on the autism spectrum; has received grant funding from internal and external agencies, including the National Institutes of Health and the Vanderbilt Institute for Clinical and Translational Research, to study the efficacy of various interventions geared toward young children with autism; and is employed by the Department of Hearing and Speech Sciences at Vanderbilt University Medical Center, which offers intervention services for autistic children via their outpatient clinics and trains clinical students in the provision of treatments delivered over the course of early childhood. MS has received fees for presenting research findings in invited talks and providing expert evidence on the effects of early intervention for young autistic children in court hearings. She previously taught courses in a program that was accredited by the Behavior Analyst Certification Board on both behavioral and NDBI early childhood interventions. The remaining authors have declared that they have no competing or potential conflicts of interest.

Ethical considerations

This project did not include human or animal subjects, and so was not subject to review by an institutional review board.

Endnotes

1. We use 'conceptual' definition to contrast with 'operational' definition. A conceptual definition describes the overall behavior construct, while an operational definition describes how instances of the behavior are tallied or measured in the context of a given study. If a study only offered an operational definition in the context of describing their measurement system, we did not count this as a conceptual definition. Ideally, studies would offer both types of definitions, so readers can assess the extent to which an operational definition adequately aligns with the conceptual definition.
2. We were unable to locate the manual for the original *Repetitive Behavior Scale* or the revised version; we were advised by the developer to consult journal publications reporting on the scale's validity.

Data availability

All data associated with this project, including a reference list of included articles, coding manual, and coding spreadsheets, are available on the Open Science Framework at the following link: https://osf.io/g57aj/?view_only=2feae004062043e6bdcd7c4cad641daf.

Correspondence

Kristen Bottema-Beutel, Lynch School of Education and Human Development, Boston College, 140 Commonwealth Ave., Chestnut Hill, MA 02467, USA; Email: kristen.bottema-beutel@bc.edu

© 2025 The Author(s). *Journal of Child Psychology and Psychiatry* published by John Wiley & Sons Ltd on behalf of Association for Child and Adolescent Mental Health.

14697610, 0, Downloaded from https://acamh.onlinelibrary.wiley.com/doi/10.1111/jcpp.14177 by Turkey Cochrane Evidence Aid, Wiley Online Library on [23/09/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

14697610, 0, Downloaded from https://acamh.onlinelibrary.wiley.com/doi/10.1111/jcpp.14177 by Turkey Cochrane Evidence Aid, Wiley Online Library on [23/09/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Key points

- Many autism interventions attempt to reduce behaviors that co-occur with autism, such as aggression, self-injury, or repetitive and restrictive behavior. Collectively, these behaviors are often referred to as 'problem behaviors'.
- We conducted a secondary systematic review of nonpharmacological intervention studies with autistic children up to age eight to examine study-, variable-, and measurement-level aspects of interventions that attempted to reduce problem behavior.
- We found that researchers conceptually defined the problem behavior they attempted to reduce only some of the time, often did not provide a rationale for why they were trying to reduce it, and most frequently measured problem behavior using off-the-shelf caregiver reports. We also found that these measures mix together assessment of behaviors that are inherently harmful with assessment of behaviors that are nonnormative but not harmful, and many of these assessments have not undergone sufficient psychometric evaluation in autistic populations.
- Our study emphasizes a need for a refined conceptualization of 'problem behavior', stronger rationale to reduce behaviors in autistic children, and validated measures of behavior that do not mix in assessment of benign or beneficial autistic traits.

References

Ala'i-Rosales, S., Cihon, J.H., Currier, T.D.R., Ferguson, J.L., Leaf, J.B., Leaf, R., … & Weinkauf, S.M. (2019). The big four: Functional assessment research informs preventative behavior analysis. *Behavior Analysis in Practice*, *12*, 222–234.

Aldabbagh, R., Glazebrook, C., Sayal, K., & Daley, D. (2024). Systematic review and meta-analysis of the effectiveness of teacher delivered interventions for externalizing behaviors. *Journal of Behavioral Education*, *33*, 233–274.

Ayvaci, A.S., Cox, A.D., & Dimopoulos, A. (2024). A quantitative systematic literature review of combination punishment literature: Progress over the last decade. *Behavior Modification*, *49*, 117–153.

Belek, B. (2019). Articulating sensory sensitivity: From bodies with autism to autistic bodies. *Medical Anthropology*, *38*, 30–43.

Beneke, M., & Rasmus, W. (1992). "Clinical Global Impressions" (ECDEU): Some critical comments. *Pharmacopsychiatry*, *25*, 171–176.

Bent, C.A., Aulich, A., Constantine, C., Fidock, E., Dwyer, P., Green, C., … & Hudry, K. (2024). Autistic and autism community perspectives on infant and family support in the first two years of life: Findings from a community consultation survey. *Autism* Advance Online Publication. https://doi.org/10.1177/13623613241262077

Blackburn, C., Tueres, M., Sandanayake, M., Roberts, J., & Sutherland, R. (2023). A systematic review of interventions for echolalia in autistic children. *International Journal of Language & Communication Disorders*, *58*, 1977–1993.

Bottema-Beutel, K. (2024). Putting autism research in social contexts. *Autism*, *28*, 1597–1601.

Bottema-Beutel, K., Kapp, S.K., Sasson, N., Gernsbacher, M.A., Natri, H., & Botha, M. (2023). Anti-ableism and scientific accuracy in autism research: A false dichotomy. *Frontiers in Psychiatry*, *14*, 1244451.

Bottema-Beutel, K., McKinnon, R., Mohiuddin, S., LaPoint, S.C., & Kim, S.Y. (2024). Problems with "problem behavior": A secondary systematic review of intervention research on transition-age autistic youth. *Autism*, *28*, 1872–1888.

Boyd, B.A., Watson, L.R., Reszka, S.S., Sideris, J., Alessandri, M., Baranek, G.T., … & Belardi, K. (2018). Efficacy of the ASAP intervention for preschoolers with ASD: A cluster randomized controlled trial. *Journal of Autism and Developmental Disorders*, *48*, 3144–3162.

Carr, E.G. (2007). Social skills that are not always social and problems that are not always problems. *Research and Practice for Persons With Severe Disabilities*, *32*, 110–111.

Case-Smith, J., Weaver, L.L., & Fristad, M.A. (2015). A systematic review of sensory processing interventions for children with autism spectrum disorders. *Autism*, *19*, 133–148.

Constantino, J.N., & Gruber, C.P. (2012). *Manual: Social Responsiveness Scale* (2nd edn). Torrance: Western Psychological Services.

Cook, J., Hull, L., Crane, L., & Mandy, W. (2021). Camouflaging in autism: A systematic review. *Clinical Psychology Review*, *89*, 102080.

Emerson, E. (2001). *Challenging behaviour: Analysis and intervention in people with severe intellectual disabilities*. Cambridge University Press.

Han, E., Scior, K., Avramides, K., & Crane, L. (2022). A systematic review on autistic people's experiences of stigma and coping strategies. *Autism Research*, *15*, 12–26.

Hanley, G.P., Iwata, B.A., & McCord, B.E. (2003). Functional analysis of problem behavior: A review. *Journal of Applied Behavior Analysis*, *36*, 147–185.

Hardesty, D.M., & Bearden, W.O. (2004). The use of expert judges in scale development: Implications for improving face validity of measures of unobservable constructs. *Journal of Business Research*, *57*, 98–107.

Harrop, C., Gulsrud, A., Shih, W., Hovsepyan, L., & Kasari, C. (2017). The impact of caregiver-mediated JASPER on child restricted and repetitive behaviors and caregiver responses. *Autism Research*, *10*, 983–992.

Hassiotis, A., & Rudra, S. (2022). Behaviours that challenge in adults with intellectual disability: Overview of assessment and management. *BJPsych Advances*, *28*, 393–400.

Higgins, J.P.T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M.J., & Welch, V.A. (Eds.). (2021). *Cochrane handbook for systematic reviews of interventions (version 6.2)*. Cochrane. Available from: https://training.cochrane.org/handbook

Higgins, W.C., Kaplan, D.M., Deschrijver, E., & Ross, R.M. (2024). Construct validity evidence reporting practices for the reading the mind in the eyes test: A systematic scoping review. *Clinical Psychology Review*, *108*, 102378.

© 2025 The Author(s). *Journal of Child Psychology and Psychiatry* published by John Wiley & Sons Ltd on behalf of Association for Child and Adolescent Mental Health.

Kanner, L. (1943). Autistic disturbances of affective contact. *The Nervous Child*, *2*, 217–250.
Kapp, S.K. (2014). A critical response to "The Kids Who Beat Autism". Thinking person's guide to autism. Available from: https://thinkingautismguide.com/2014/08/a-critical-response-to-kids-who-beat.html
Kapp, S.K., Steward, R., Crane, L., Elliott, D., Elphick, C., Pellicano, E., & Russell, G. (2019). 'People should be allowed to do what they like': Autistic adults' views and experiences of stimming. *Autism*, *23*, 1782–1792.
Kim, S.Y., & Bottema-Beutel, K. (2019). Negotiation of individual and collective identities in the online discourse of autistic adults. *Autism in Adulthood*, *1*, 69–78.
Lam, K.S., & Aman, M.G. (2007). The Repetitive Behavior Scale-Revised: Independent validation in individuals with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *37*, 855–866.
Leadbitter, K., Aldred, C., McConachie, H., Le Couteur, A., Kapadia, D., Charman, T., … & Green, J. (2018). The Autism Family Experience Questionnaire (AFEQ): An ecologically-valid, parent-nominated measure of family experience, quality of life and prioritised outcomes for early intervention. *Journal of Autism and Developmental Disorders*, *48*, 1052–1062.
Leadbitter, K., Buckle, K.L., Ellis, C., & Dekker, M. (2021). Autistic self-advocacy and the neurodiversity movement: Implications for autism early intervention research and practice. *Frontiers in Psychology*, *12*, 635690.
Leader, G., Tuohy, E., Chen, J.L., Mannion, A., & Gilroy, S.P. (2020). Feeding problems, gastrointestinal symptoms, challenging behavior and sensory issues in children and adolescents with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *50*, 1401–1410.
Lord, C., & Rutter, M. (2012). *ADOS-2: Autism Diagnostic Observation Schedule* (2nd edn). Torrance: Western Psychological Services.
Lovaas, O.I., & Smith, T. (1989). A comprehensive behavioral theory of autistic children: Paradigm for research and treatment. *Journal of Behavior Therapy and Experimental Psychiatry*, *20*, 17–29.
Maddox, B.B., Trubanova, A., & White, S.W. (2017). Untended wounds: Non-suicidal self-injury in adults with autism spectrum disorder. *Autism*, *21*, 412–422.
Medeiros, K., Mazurek, M.O., & Kanne, S. (2017). Investigating the factor structure of the Child Behavior Checklist in a large sample of children with autism spectrum disorder. *Research in Autism Spectrum Disorders*, *40*, 24–40.
National Collaborating Center for Mental Health (UK). (2015). *Challenging behavior and learning disabilities: Prevention and intervention for people with learning disabilities whose behavior challenges*. London: National Collaborating Center for Mental Health (UK). Available from: https://www.ncbi.nlm.nih.gov/books/NBK355392
O'Haire, M.E. (2013). Animal-assisted intervention for autism spectrum disorder: A systematic literature review. *Journal of Autism and Developmental Disorders*, *43*, 1606–1622.
Phung, J., Penner, M., Pirlot, C., & Welch, C. (2021). What I wish you knew: Insights on burnout, inertia, meltdown, and shutdown from autistic youth. *Frontiers in Psychology*, *12*, 741421.
Pruccoli, J., Spadoni, C., Orsenigo, A., & Parmeggiani, A. (2021). Should echolalia be considered a phonic stereotypy? A narrative review. *Brain Sciences*, *11*, 862.
Rapp, J.T., & Vollmer, T.R. (2005). Stereotypy I: A review of behavioral assessment and treatment. *Research in Developmental Disabilities*, *26*, 527–547.
Raymaker, D.M., Teo, A.R., Steckler, N.A., Lentz, B., Scharer, M., Delos Santos, A., … & Nicolaidis, C. (2020). "Having all of your internal resources exhausted beyond measure and being left with no clean-up crew": Defining autistic burnout. *Autism in Adulthood*, *2*, 132–143.
Sandbank, M., Bottema-Beutel, K., LaPoint, S.C., Feldman, J.I., Barrett, D.J., Caldwell, N., … & Woynaroski, T. (2023). Autism intervention meta-analysis of early childhood studies (project AIM): Updated systematic review and secondary analysis. *BMJ*, *383*, e076733.
Sandbank, M., Chow, J., Bottema-Beutel, K., & Woynaroski, T. (2021). Evaluating evidence-based practice in light of the boundedness and proximity of outcomes: Capturing the scope of change. *Autism Research*, *14*, 1536–1542.
Schaber, A. (2014). Ask an autistic #20- What are autistic shutdowns? *Youtube*. Available from: https://www.youtube.com/watch?v=3WIiL8vBjq0&t=1s [last accessed 29 July 2024].
Simplican, S.C. (2019). Behaviors that challenge disability studies. *Disability and Society*, *34*, 1379–1398.
Sivathasan, S., Eldeeb, S., Northrup, J.B., Antezana, L., Ionadi, A., Wakschlag, L.S., & Mazefsky, C.A. (2024). Early childhood aggression in autistic and non-autistic preschoolers: Prevalence, topography, and relationship to emotional reactivity. *JAACAP Open*, *2*, 112–125.
Skivington, K., Matthews, L., Simpson, S.A., Craig, P., Baird, J., Blazeby, J.M., … & Moore, L. (2021). A new framework for developing and evaluating complex interventions: Update of Medical Research Council guidance. *BMJ*, *374*, n2061.
Sterponi, L., & Shankey, J. (2014). Rethinking echolalia: Repetition as interactional resource in the communication of a child with autism. *Journal of Child Language*, *41*, 275–304.
Tarbox, J., Wilke, A.E., Najdowski, A.C., Findel-Pyles, R.S., Balasanyan, S., Caveney, A.C., … & Tia, B. (2009). Comparing indirect, descriptive, and experimental functional assessments of challenging behavior in children with autism. *Journal of Developmental and Physical Disabilities*, *21*, 493–514.
Turner-Brown, L., Hume, K., Boyd, B.A., & Kainz, K. (2019). Preliminary efficacy of family implemented TEACCH for toddlers: Effects on parents and their toddlers with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *49*, 2685–2698.
Wigham, S., & McConachie, H. (2014). Systematic review of the properties of tools used to measure outcomes in anxiety intervention studies for children with autism spectrum disorders. *PLoS One*, *9*, e85268.
Wilkinson, E., Farmer, C., Kleiman, E., & Bal, V.H. (2023). Factor structure of the VABS-3 comprehensive parent/caregiver form in autistic individuals: Poor fit of three-factor and unidimensional models. *Autism*, *28*, 616–626.

Accepted for publication: 12 March 2025

© 2025 The Author(s). *Journal of Child Psychology and Psychiatry* published by John Wiley & Sons Ltd on behalf of Association for Child and Adolescent Mental Health.

14697610, 0, Downloaded from https://acamh.onlinelibrary.wiley.com/doi/10.1111/jcpp.14177 by Turkey Cochrane Evidence Aid, Wiley Online Library on [23/09/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License