Screening, Assessment, and Measurement of Challenging Behavior in Autism Spectrum Disorder

William E. Sullivan, PhD
Golisano Center for Special Needs
SUNY Upstate Medical University

Amanda N. Zangrillo, PsyD, BCBA-D
University of Nebraska Medical Center

Luther G. Kalb, PhD, MHS
Kennedy Krieger Institute
Johns Hopkins Bloomberg School of Public Health

Carla A. Mazefsky, PhD
University of Pittsburgh
Acknowledgments

The Autism Speaks Thought Leadership Summit on Challenging Behaviors, held in December 2020, convened leaders in autism care and research across North America. The Summit aimed to characterize the landscape of services and supports for people with autism with challenging behavior and acted as a catalyst for innovations in programs and policies to improve systems of care for this population. Six workgroups were formed, utilizing Summit participants, to develop recommendations and priorities related to both practice and public policy.

The following document covers screening, assessment, and measurement of challenging behavior, a need identified by Summit leaders and participants, to better understand what is currently known about outcomes for autistic individuals with severe challenging behaviors, and which practices, programs, and policies have demonstrated impacts; what existing or new mechanism can help in scaling up such efforts; and what potential opportunities exist to enhance capacities in systems supporting services to aid in scaling-up or improving quality of implementation in this topical area.

Special thanks to Autism Speaks for driving and supporting this initiative, to Jordan DeBrine and Jamie Peven for contributing to this effort, and to SUNY Upstate Medical University, the University of Nebraska Medical Center, John Hopkins Bloomberg School of Public Health, and the University of Pittsburgh for lending additional support and resources. Additional thanks to Thought Summit leaders Drs. Matthew Siegel, Henry Roane, Eric Butter, and Donna Murray, as well as Jackie Perlmeter, for their leadership, feedback, and guidance throughout the process.
How to Cite this Document


{Autism Speaks URL for the document}
SCREENING AND ASSESSMENT

Outline

- Introduction

- Prevalence of Challenging Behavior among Children with ASD
  - Introduction
  - Methods
  - Results
    - Aggression
    - Self-Injurious Behavior
    - Elopement
    - Mental Health Crises in ASD
  - Summary

- Screeners and Standardized Assessments
  - Informant-Based Screeners
  - Direct Screeners
  - Standardized Measures on Challenging Behaviors
  - Psychiatric Assessment
  - Summary

- Functional Behavior Assessment
  - Indirect Assessment
  - Direct Measurement
    - Continuous Measurement
    - Discontinuous Measurement
  - Descriptive Assessment
  - Functional Analysis
  - Summary

- Multi-Tiered Systems of Support
  - Primary-Care Pathway
  - Progress Monitoring
  - Summary

- Conclusion
SCREENING AND ASSESSMENT

Screening, Assessment, and Measurement of Challenging Behavior in Autism Spectrum Disorder

Introduction

Children commonly display challenging behaviors in early childhood as they learn to navigate their social worlds and develop language (Fletcher, 2011). Challenging behaviors (e.g., tantrums, self-injury), while considered a normal part of development, tend to decrease between the ages of 3 and 5 years old given development of positive social behaviors (e.g., language skills and emotional regulation; Underwood, 2003). For some children, challenging behavior persists beyond this typical developmental window. These behaviors begin to gain the attention of caregivers, physicians, educators, and other key stakeholders when they produce emotional or physical harm to the individual or others, and impact the individual’s participation in home, school, or community settings. Individuals with autism spectrum disorder (ASD) and other developmental disabilities, demonstrate a higher prevalence of challenging behavior relative to same-aged peers without documented diagnoses (Fahmie et al., 2020; Hill et al., 2014; Kanne & Mazurek, 2011, Schroeder et al., 2014). This is of particular importance as individuals displaying challenging behaviors experience increased risk for persistent stress, abuse, decreased access to much needed supports and services, as well as caregiver burnout, and lower quality of life (Fitzpatrick et al., 2016).

There are several theories that have attempted to explain the etiology of challenging behavior displayed by individuals with ASD. Behavior equivalents theory (Emerson, 2001) suggests that challenging behavior may be an alternative manifestation of psychopathology. For example, social situations may elicit feelings of anxiety that manifest as challenging behavior for individuals with ASD. Neurobiological models implicate brain dysfunction as a cause for challenging behavior. Frontal cortex dysfunction, leading to poor inhibitory control, and
SCREENING AND ASSESSMENT

increased activation of the amygdala and hypothalamus have been suggested to contribute to aggression (e.g., Siever, 2008); whereas impairment of the basal ganglia and fronto-striatal circuits have been linked to self-injurious behavior (Bodfish & Lewis, 2002; Turner & Lewis, 2002). A variety of underlying genetic factors that produce specific behavioral phenotypes are also associated with challenging behavior (e.g., Oliver et al., 2013). Furthermore, difficulty regulating emotions has been shown to contribute to challenging behavior when environmental demands exceed self-regulatory capacity. Individuals with ASD are four times more likely than the general population to exceed clinical cut-offs for impairing emotion dysregulation (Conner et al., 2021), which has been specifically associated with aggression (Mazefsky et al., 2018, Conner et al., 2020; Northrup et al., in press).

The most prominent theory explaining the cause of challenging behavior, however, suggests that challenging behavior is learned through the process of operant conditioning (Skinner, 1938). In this model, it is purported that challenging behavior is selected by the environmental situations that evoke it and the consequences that maintain it. For example, if an individual with ASD engages in challenging behavior when their caregiver’s attention is diverted, and when the challenging behavior occurs attention is provided, the individual may learn to engage in challenging behavior to obtain caregiver attention. That is, challenging behavior may be reinforced by producing or being associated with a favorable consequence, and thus, likely to occur again in the future under similar conditions.

Taken together, the factors that contribute to challenging behavior in ASD are complex. It is plausible that many of the above factors, in combination, contribute to the occurrence and maintenance of challenging behavior. For example, genetics and patterns of neural reactivity may contribute to underlying irritability and poor emotion regulation that is understood as
SCREENING AND ASSESSMENT

manifesting as challenging behavior when an individual is confronted with an aversive situation. Then, when challenging behavior results in the removal of that aversive situation, the individual may learn that challenging behavior is an effective way to communicate their needs and continue to engage in challenging behavior when confronted with similar situations.

A variety of evidence-based approaches to assessment and treatment have been developed and shown to effectively reduce challenging behavior displayed by individuals with ASD (refer to Evidence Based Practices document). The dosage, complexity, and modality of these interventions vary widely and will ultimately depend on the needs of the individual and the capacity of the local/regional support system. An individual who engages in severe forms of self-injurious behavior (SIB), for example, may require more intense intervention than a child who engages in mild disruptive behavior. Moreover, individuals with ASD who engage in challenging behavior and their families often solicit help from multiple support systems. The child’s parents may report these challenging behaviors to their pediatrician, the teacher may make a referral to their school psychologist, or community mental-health providers may offer family support. Across these systems, however, there is not a recognized and widely disseminated systematic framework for: (a) screening for challenging behavior in individuals with ASD, (b) connecting them with appropriate treatment services, and (c) surveillance of these individuals over time to promote behavioral health across the lifespan.

Overall, the goal of this document is to provide general recommendations regarding the identification and assessment of challenging behavior in ASD based on the existing literature, provide directions for future research, and highlight areas for advocacy to promote the well-being of those affected by challenging behavior in the ASD population. In the following sections we first provide a review of the literature concerning the prevalence of challenging behavior and
SCREENING AND ASSESSMENT

mental-health crises among children with ASD. Next, we describe the broad and complex process of screening, standardized assessment, and function-based assessment of challenging behavior in ASD. Finally, a multi-tiered system of support (MTSS; Sugai & Horner, 2002) is discussed as an example of a framework in which challenging behavior displayed by individuals with ASD can be assessed, linked to appropriate treatments, and monitored over time to promote wellness.

Prevalence of Challenging Behavior among Children with ASD

The prevalence of psychiatric and behavioral disorders is elevated among youth with ASD. In community samples, 70% of youth with ASD have at least one psychiatric disorder, a prevalence that is 3.5 times greater than in the non-ASD population (Simonoff et al., 2008). The prevalence of comorbid psychiatric disorders is even higher in clinical samples with 95% and 74% of youth with ASD exhibiting three and five disorders (e.g., attention deficit hyperactivity disorder, anxiety disorders), respectively (Joshi et al., 2010). The pressing need to manage these psychiatric and behavioral symptoms has resulted in high rates of polypharmacy, emergency room visits, and inpatient psychiatric hospitalization among youth with ASD (Jobski et al., 2016; Kalb et al., 2012; Mandell, 2008). There is even evidence to suggest youth with ASD are four times more likely to visit the emergency room for mental-health purposes, relative to their peers without ASD (Liu et al., 2017).

While data on diagnoses, treatments, and service use all point to increased prevalence of challenging behaviors, data are lacking on the population prevalence of specific challenging behaviors. Understanding the prevalence of specific challenging behaviors is critical as they may not fit neatly into diagnostic categories (e.g., elopement). The purpose of this section is to provide a comprehensive, up-to-date review of the literature concerning the prevalence of
SCREENING AND ASSESSMENT

challenging behavior among children with ASD. Three behaviors were evaluated: 1) aggression; 2) SIB; and 3) elopement. These behaviors were selected as they are among the most common and impairing externalizing problems for children with ASD. Note this section is not intended to meet the criteria of systematic review. Rather, the findings provide current information on the prevalence of these challenging behaviors to set the stage for the following discussions.

Methods

The literature search was conducted using PsychInfo, Google Scholar, and PubMed databases. Search terms were performed by including variations of a particular behavior with autism terms. For instance, the search terms used for aggression were “(aggression OR aggressive behavior OR aggressiveness) AND (autism OR ASD OR autism spectrum disorder).” For SIB, the search terms were “(self-injurious behavior OR self-injury OR self-harm) AND (autism OR ASD OR autism spectrum disorder).” Finally, the search terms for elopement were “(elopement OR elope OR wandering) AND (autism OR ASD OR autism spectrum disorder).”

Each study was screened for inclusion using several additional criteria. Inclusion criteria were: 1) the study must be published within the last 10 years; 2) the study has a sample size of at least 100 participants; and 3) the participants must be predominantly less than 18 years of age. Studies focusing on adults with ASD were excluded given their relative scarcity, differential approaches to informant and measurement, and small sample sizes.

Once an article met inclusion criteria, a custom database housed information about each study. The database captured: 1) authorship; 2) year of publication; 3) sample size; 4) setting; 5) measurement; and 6) prevalence. Setting was classified as single-site clinical study, multi-site clinical study, population-based (i.e., a community-based study or a national survey), or meta-analysis. Measurement details, concerning the actual tool used to assess the behavior and
relevant cutoffs, were captured. Prevalence reflected the proportion of the overall sample that met criteria for the behavior. Details on how prevalence was defined is provided.

<table>
<thead>
<tr>
<th>Author</th>
<th>Year of Publication</th>
<th>Sample Size</th>
<th>Setting</th>
<th>Measurement</th>
<th>Prevalence</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kanne and Mazurek</td>
<td>2011</td>
<td>1380</td>
<td>Population-based study</td>
<td>ADI-R Aggression Items</td>
<td>68% towards parents 49% towards non-caregiver</td>
<td>Yes/No No Aggression – score of 0 on aggression questions Definite Aggression – score of 2 or 3 aggression questions</td>
</tr>
<tr>
<td>McTiernan et al.</td>
<td>2011</td>
<td>174</td>
<td>Population-based study</td>
<td>BPI-01</td>
<td>56.3%</td>
<td>Aggressive Subscale, 1 or 11 items rated weekly or more</td>
</tr>
<tr>
<td>Maskey et al.</td>
<td>2012</td>
<td>843</td>
<td>Population-based study</td>
<td>Parent Questionnaire developed by study authors</td>
<td>21.8%</td>
<td>Yes/No Parent Report Question “Aggression towards other people”</td>
</tr>
<tr>
<td>Mayes et al.</td>
<td>2012</td>
<td>435</td>
<td>Single-site Clinical study</td>
<td>PBS – 8 aggression items</td>
<td>16.6%</td>
<td>Raw Scores on subscale items or Maternal Ratings of behavior problem (No = 0, Yes = 1, 2, or 3)</td>
</tr>
<tr>
<td>Medeiros et al.</td>
<td>2012</td>
<td>221</td>
<td>Single-Site Clinical Study</td>
<td>BISCA Part 3 for Challenging Behavior</td>
<td>78.5%</td>
<td>Yes/No No = item score of 0 Yes = item score of 1 or 2</td>
</tr>
<tr>
<td>Mazurek et al.</td>
<td>2013</td>
<td>1584</td>
<td>Multi-site Clinical Study</td>
<td>Autism Treatment Network (ATN) – single item from ATN Parent Survey</td>
<td>53%</td>
<td>Yes/No Is child currently demonstrating physical aggression?</td>
</tr>
<tr>
<td>Hill et al.</td>
<td>2014</td>
<td>400</td>
<td>Single-Site Clinical Study</td>
<td>CBCL ABC</td>
<td>25%</td>
<td>T-score &gt; 70</td>
</tr>
<tr>
<td>Farmer et al.</td>
<td>2015</td>
<td>414</td>
<td>Multi-site Clinical Study</td>
<td>CBCL Aggression Subscale</td>
<td>19%</td>
<td>T-score &gt; 70</td>
</tr>
</tbody>
</table>

Note. Autism Diagnostic Interview – Revised (ADI-R); Baby and Infant Screening for Child with Autism (BISCA); Behavior Problem Inventory (BPI-01); Child Behavior Checklist (CBCL); Pediatric Behavior Scale (PBS)
SCREENING AND ASSESSMENT

Results

Shown in Table 1, eight studies met inclusion criteria for Aggression. These studies included a total of 5,451 participants. Prevalence of aggression ranged from 17% to 68%, with a pooled prevalence of 42%. The settings in which these studies were conducted included population-based (37%), single-site clinical (37%), and multi-site clinical (25%). Two of the eight studies used the Aggressive Behavior subscale from the Child Behavior Checklist (CBCL). The remaining six studies used a variety of different measures which included: 1) the Autism Diagnostic Interview – Revised (ADI-R); 2) the Behavior Problem Inventory (BPI-01); 3) the Pediatric Behavior Scale (PBS); 4) the Baby and Infant Screening for Child with Autism (BISCA); and 5) custom aggression items from study specific questionnaires. Of the seven different measures used to assess aggression, 29% were single-item measures, whereas the majority (71%) of the measures used consisted of multi-item assessments. Given aggression was broadly defined, using both single items and subscales, the results are heterogeneous. They reflect a variety of aggressive behaviors, including aggression (e.g., hitting, kicking), verbal aggression (e.g., yelling, inappropriate language), and aggression to property and/or objects. As such, these findings should be interpreted accordingly.

Eleven studies were included in the SIB literature review; see Table 2 for details. This constituted a total of 10,968 participants and one meta-analysis, of which 14,379 participants were included in the meta-analysis. For the ten unique studies, the prevalence of SIB ranged from 14% to 67%, with a pooled prevalence of 46%. The meta-analysis provided a very similar prevalence estimate at 42%. The settings in which these studies were conducted primarily included population-based (45%) and multi-site clinical studies (36%), with one single-site clinical study and one meta-analysis. Excluding the meta-analysis, 20% of the studies used the
SCREENING AND ASSESSMENT

SIB subscale of the BPI-01, and 30% used the SIB subscale of the Repetitive Behavior Scale-Revised (RBS-R). Half of the remaining sample used a variety of different measures, including the ADI-R, the Aberrant Behavior Checklist (ABC), the Challenging Behavior Questionnaire (CBQ), the Social and Communication Questionnaire (SCQ), and custom items. Of the seven different measures used to assess SIB, 29% were single item measures with the majority (71%) consisting of multi-item scales.
<table>
<thead>
<tr>
<th>Author</th>
<th>Year of Publication</th>
<th>Sample Size</th>
<th>Setting</th>
<th>Measurement</th>
<th>Prevalence</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>McTiernan et al.</td>
<td>2011</td>
<td>174</td>
<td>Population-based study</td>
<td>BPI-01</td>
<td>48.9%</td>
<td>SIB Subscale, 1 or 15 items rated weekly or more</td>
</tr>
<tr>
<td>Duerden et al.</td>
<td>2012</td>
<td>241</td>
<td>Multi-site Clinical Study</td>
<td>ADI-R (241 parent responses)/RBS-R: 171 Parent responses</td>
<td>ADI-R: 52.3%/RBS-R: 64.9%</td>
<td>Severity (S) Likert scale: 0 - 3 Frequency (F) Likert scale: 0 - 4 SIB-S and SIB-F = 0 – No SIB Raw score over 3 in SIB-F. = High F. Raw score over 2 in SIB-S. = High S.</td>
</tr>
<tr>
<td>Maskey et al.</td>
<td>2012</td>
<td>843</td>
<td>Population-based study</td>
<td>Parent Questionnaire developed by study authors</td>
<td>14%</td>
<td>Yes/No Parent Report Question “Injury to Self” Yes – Sometimes, Frequently, Past Only No - Never</td>
</tr>
<tr>
<td>Rattaz et al.</td>
<td>2015</td>
<td>152</td>
<td>Multi-site Clinical study</td>
<td>ABC</td>
<td>35.8%</td>
<td>Likert scale: 0 (behavior not a problem) to 3 (behavior significant problem) No SIB - total score of 0 on 3 items Low SIB - total score of 1 or 2 on 3 items High SIB - total score ≥ 3 on 3 items</td>
</tr>
<tr>
<td>Siegel et al.</td>
<td>2015</td>
<td>147</td>
<td>Multi-Site Clinical Study</td>
<td>RBS-R Self-injury subscale</td>
<td>26.5%</td>
<td>Yes/No Parent Report SIB defined as the presence of at least daily attempts at self-injury, as determined by the unit psychologist or board-certified behavior analyst.</td>
</tr>
<tr>
<td>Soke et al.</td>
<td>2016</td>
<td>8065</td>
<td>Population-based study</td>
<td>Autism and Developmental Disabilities Monitoring Network</td>
<td>27.7%</td>
<td>Yes/No (Present or Not Present in the child’s available records of any behaviors that were considered as SIB by the ADDM clinician who reviewed child's records to determine if child met the ADDM Network case definition)</td>
</tr>
<tr>
<td>Richards et al.</td>
<td>2017</td>
<td>208</td>
<td>Population-based study</td>
<td>CBQ – 8 items about topography of SIB</td>
<td>45.7%</td>
<td>Yes/No: No = 0, Yes = 1,2,3, or 4</td>
</tr>
<tr>
<td>Handen et al.</td>
<td>2018</td>
<td>302</td>
<td>Multi-site Clinical study</td>
<td>RBS-R – 8 SIB item subscale</td>
<td>67.5% - home and/or hospital 49.2% home only 24.8% home and hospital</td>
<td>Low/No SIB - score of less than 2 for any form of SIB on parent report RBS-R Home SIB - score of 2 or more on at least one SIB item of RBSR by parent Hospital SIB - No parent report of 2 or more on SIB items, but observed engaging in SIB daily while inpatient Home and Hospital SIB - Both parent report of SIB and daily observation of SIB inpatient</td>
</tr>
</tbody>
</table>
**SCREENING AND ASSESSMENT**

<table>
<thead>
<tr>
<th>Study</th>
<th>Year</th>
<th>Participants</th>
<th>Study Type</th>
<th>Parent Report</th>
<th>Prevalence Estimate</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>Soke et al.</td>
<td>2018</td>
<td>691</td>
<td>Population-based study</td>
<td>SIB based on question in SCQ</td>
<td>29.4%</td>
<td>Yes/No Parent report to question: &quot;Does the child ever injure her/himself deliberately, such as by biting her/his arm or banging her/his head&quot;</td>
</tr>
<tr>
<td>Flowers et al.</td>
<td>2020</td>
<td>145</td>
<td>Single-site Clinical Study</td>
<td>BPI-01 – Short Form</td>
<td>50% - High F./High S. 50% - High F./High S. 11% - Low F./Low S. 11% - Low F./Low S. 10% - Low F./High S. 10% - Low F./High S. 6% High F./Low S. 6% High F./Low S.</td>
<td>Severity (S) Likert scale: 0 - 3 Frequency (F) Likert scale: 0 - 4 SIB-S and SIB-F = 0 – No SIB Raw score over 3 in Freq. = High Freq. Raw score over 2 in Sev. = High Severity</td>
</tr>
<tr>
<td>Steenfeldt-</td>
<td>2020</td>
<td>14,379</td>
<td>Meta-analysis</td>
<td></td>
<td>Yes, 42%</td>
<td>Multiple measures</td>
</tr>
<tr>
<td>Kristensen et al.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* Aberrant Behavior Checklist (ABC); Autism Diagnostic Interview – Revised (ADI-R); Behavior Problem Inventory (BPI-01); Challenging Behavior Questionnaire (CBQ); Repetitive Behavior Scale-Revised (RBS-R); Social and Communication Questionnaire (SCQ).

Shown in Table 3, a total of eight studies met inclusion criteria for *Elopement*, reporting on 9,398 participants. The pooled prevalence estimate of elopement was 48%, with estimates ranging from 27% to 68%. Almost all the studies (87%) were population-based, with one single site study. Six of the studies used a custom item/questionnaire, with the remaining using the CBCL or the Children’s Social Behavior Questionnaire (C-SBQ).
<table>
<thead>
<tr>
<th>Author</th>
<th>Year of Publication</th>
<th>Sample Size</th>
<th>Setting</th>
<th>Measurement</th>
<th>Prevalence</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anderson et al.</td>
<td>2012</td>
<td>1218</td>
<td>Population-based study</td>
<td>Elopement Questionnaire (developed by study authors)</td>
<td>49%</td>
<td>Yes/No (Ever vs. Never and Missing (for children were gone long enough to cause concern) vs. non-Missing)</td>
</tr>
<tr>
<td>Kiely et al.</td>
<td>2016</td>
<td>1416</td>
<td>Population-based study</td>
<td>C-SBQ</td>
<td>26.7%</td>
<td>Yes/No/Don't Know/Refuse Prevalence calculated as percentage of those responding with &quot;Yes&quot;</td>
</tr>
<tr>
<td>Rice et al.</td>
<td>2016</td>
<td>1420</td>
<td>Population-based study</td>
<td>National Survey of Children with Special Health Care Needs (NS-CSHCN)</td>
<td>37.7%</td>
<td>Yes/No to Question: &quot;Within the past year, has [Child] wandered off or became lost from a. your home? b. someone else’s home such as a relative, friend, neighbor, or babysitter? c. school, day care, or summer camp? d. a store, restaurant, playground, campsite, or any other public place?&quot;</td>
</tr>
<tr>
<td>Barnard-Brak et al.</td>
<td>2016</td>
<td>1744</td>
<td>Population-based study</td>
<td>NS-CSHSN Elopement Questions</td>
<td>28%</td>
<td>Yes/No to Question: &quot;Within the past year, has [Child] wandered off or became lost from a. your home? b. someone else’s home such as a relative, friend, neighbor, or babysitter? c. school, day care, or summer camp? d. a store, restaurant, playground, campsite, or any other public place?&quot;</td>
</tr>
<tr>
<td>McLaughlin et al.</td>
<td>2018</td>
<td>1454</td>
<td>Population-based study</td>
<td>Anonymous Online Questionnaire</td>
<td>68%</td>
<td>Yes/No Parent Report “Has your child ever wandered from adult supervision?”</td>
</tr>
<tr>
<td>Andersen et al.</td>
<td>2019</td>
<td>526</td>
<td>Population-based study</td>
<td>Elopement Prevention Questionnaire (developed by study authors)</td>
<td>49%</td>
<td>Participants were asked to select from one of seven options regarding the frequency that their child/dependent tried to leave safe spaces and/or the supervision of caregivers during the past year, with options ranging from less than once a month to many times daily</td>
</tr>
<tr>
<td>Pereira-Smith et al.</td>
<td>2019</td>
<td>394</td>
<td>Single-site clinical study</td>
<td>Questionnaire developed by study authors</td>
<td>68%</td>
<td>Yes/No (Likert Scale: No: Never (Zero), Yes: Occasionally (1-2 times), Frequently (3-5 times), Very Frequently (5+ times)</td>
</tr>
<tr>
<td>Wiggins et al.</td>
<td>2020</td>
<td>1196</td>
<td>Population-based study</td>
<td>CBCL</td>
<td>60.4%</td>
<td>Yes/No No - Parent response = not true Yes - Parent response = Sometimes true or often true</td>
</tr>
</tbody>
</table>

*Note. Child Behavior Checklist (CBCL); Children’s Social Behavior Questionnaire (C-SBQ); National Survey of Children with Special Health Care Needs (NS-CSHCN)*
Beyond the challenging behaviors described above, the core features of ASD (i.e., social deficits, communication delays, and the presence of restricted and repetitive behaviors) may themselves be an intrinsic diathesis for crisis. For instance, complex interactions between ASD symptoms, cognitive features of ASD (e.g., intellectual disability, Newschaffer et al., 2007), and environmental stressors (e.g., bullying; Zablotsky et al., 2014) may place these children at high risk for crisis compared to their neurotypical peers. The term “mental health crisis” is analogous to a psychiatric emergency, which according to the American Psychiatric Association (APA), includes two key components: (a) an acute psychiatric event that requires immediate intervention and (b) the lack of perceived resources to manage the event (Allen et al., 2002). In ASD, the term mental health crisis may be more befitting than a psychiatric emergency since behavior problems associated with crisis in this population often do not neatly fit into a psychiatric diagnosis. For example, elopement, which is not a psychiatric disorder and refers to a child wandering or running away, is a common problem in ASD and can lead to dangerous outcomes (Anderson et al., 2012).

Although the concept of crisis has been well established, only a handful of measures that capture this construct exist. Currently available instruments include the Psychiatric Emergency Service Interview (Perlmutter & Jones, 1985), the Crisis Rating Scale (Bengelsdorf et al., 1984), the Color-Risk Psychiatric Triage Scale (Molina-Lopez et al., 2016), the Crisis Triage Rating Scale (Bengelsdorf et al., 1984), the Triage Assessment Form (Hamm et al., 2010), and the Crisis Risk and Adaptive Functioning Tool (Stokoe, 2012). The main drawbacks of these scales are their narrow focus on suicidality as the precipitating psychiatric event as well as their reliance on clinician administration, which limits the measure’s use in epidemiologic research.
SCREENING AND ASSESSMENT
due to the costs of conducting clinical assessments in large populations. Most importantly, none
of these measures were designed for youth, in general, or specific populations who may be at
greatest risk for crisis, namely those with ASD.

Recently, Weiss and Lunsky explored the measurement of crisis among families raising
a child with ASD (Weiss & Lunsky, 2011; Weiss et al., 2014). Their qualitative work suggested
that parent’s conceptualized crisis across four themes: (a) the child’s behavioral problems
and difficulty with service providers, (b) the deleterious effects of crisis on the family, (c)
frequent use of emergency services for crises, and (d) the parents’ need for social and
professional support to manage crisis-related events. These themes informed the development of
the Brief Family Distress Scale (BFDS; Weiss et al., 2014), a single item rating that considers
the global or overall state of the family as it relates to crisis. Their data show that the BFDS is
positively associated with a host of adverse outcomes including negative life events, financial
problems, poor quality of life, and the child’s problem behavior (Weiss et al., 2014). While the
BFDS differs from previous crisis measures because it focuses on availability of family
resources versus psychopathology (e.g., suicidality) as the precipitant of crisis, this item is likely
valuable in identifying families who could use a referral for social work, family navigation
services and/or close psychosocial monitoring by the primary care provider (PCP).

To overcome the shortage of assessment tools available to evaluate whether youth with
ASD are at risk for a mental health crisis, Kalb et al. (2017) developed the Mental Health Crisis
Scale (MCAS). This publicly available, caregiver-report measure takes about 10 minutes to
complete. There are three sections of the MCAS with the first two sections identifying the
severity of behaviors while items from section three are summed for a total score. A cutoff of 16
places the child at risk for a crisis. This cutoff is highly accurate in relation to clinicians’
SCREENING AND ASSESSMENT
determination of crisis (ROC=.86; 83% Sensitivity; 86% Specificity; 88% Correctly Classified).
The original MCAS development study also demonstrated strong alignment with clinician
determination of crisis (Kalb et al., 2018). Beyond criterion validity, the MCAS has
demonstrated strong psychometric characteristics. This includes construct (via factor-analytic
methods), concurrent (correlations with other related constructs, like the BFDS and parental
stress), and ecological validity (association with previous psychiatric hospitalization; Kalb et al.,
2018). The measure is also reliable, as measured by internal consistency (Cronbach’s alpha = .88).

The MCAS has been used to measure the prevalence of mental health crisis in children
and young adults with ASD in two studies, one of which is published (Vasa et al., 2020) and the
other in development. In the first study, the prevalence of crisis across a sample of youth and
young adults with ASD at 32% (Vasa et al., 2020). These data were gathered from an online
sample of N = 462 community youth and young adults with ASD (not clinically referred).
Younger age, increased parental depressive symptoms, and lower family quality of life were
significantly correlated with higher crisis scores (Vasa et al., 2020).

The elevated estimate of crisis among youth with ASD suggests a need for substantial
mental-health support. Unfortunately, results from a national study of child and adolescent
psychiatrists found these providers lacked access to specialized resources, such as social workers
and psychiatric crisis evaluation centers, needed to assess and treat mental-health crises in youth
with ASD (Kalb et al., 2016). That finding is concurrent with parental-report in terms of
difficulty accessing high-quality mental-health care for their child (Brookman-Frazee et al.,
2010; Chiri et al., 2012). Psychiatrists also expressed concerns about the ability of emergency
department (ED) professionals and emergency responders to manage mental-health crises among
SCREENING AND ASSESSMENT

Youth with ASD in a safe and developmentally appropriate fashion (Kalb et al., 2016). These findings, coupled with a recent national study showing only half of mental-health providers will treat youth with ASD, suggest greater outpatient services are needed for those with ASD (Cantor et al., 2021).

Summary

Challenging behavior among children with ASD is highly prevalent. After reviewing studies that covered almost 25,000 children, greater than 4 in 10 engaged in at least one challenging behavior. Interestingly, the prevalence of each behavior was quite similar. Moreover, mental-health crises in youth and young adults with ASD have been reported to be common. It should be noted that a preponderance of these studies were based in clinical settings (i.e., many of these children were likely referred for these or related challenges). Nevertheless, population-based studies, that present less issues with selection bias, often produced similar estimates.

Another important finding is the measurement variability within and across constructs and studies. This heterogeneity makes cross-study comparisons difficult. This is particularly true when attempting to: a) synthesize the literature regarding severity and b) identify risk factors for challenging behaviors. When parents serve as the informant, use of standardized, normative measures such as the CBCL is recommended. Normative measures naturally provide comparison groups to understand severity/level of impairment, through the use of T-scores.

Screeners and Standardized Assessments

Given the prevalence of challenging behavior in the ASD population, and the potential negative health outcomes associated with its occurrence (e.g., persistent stress, lower quality of life; Fitzpatrick et al., 2016), early identification and targeted assessment is needed. Screeners and informant-based measures of challenging behavior can be helpful for identifying problems,
SCREENING AND ASSESSMENT
determining the degree of impairment, and in some cases (when shown to be reliable and
sensitive to change), monitoring progress. In the following sections we detail screening and
standardized assessment procedures used to assess challenging behavior in the ASD population.
We highlight common informant-based and direct screeners, standardized questionnaires that
focus on challenging behaviors, and psychiatric assessments. This is not a systematic review of
available questionnaires and their psychometrics; rather it provides an overview of commonly
used measures and highlights their strengths and weaknesses.

Informant-based Screeners

One class of standardized questionnaires includes broad-based screeners that tap various
forms of emotional and behavioral difficulties. Benefits of these measures is that they assess a
wide range of problems, provide standardized norms for comparison, and have versions that
cover the full lifespan. A drawback is that they are generally all copyrighted; thus, both
administration and scoring forms (or online uses) must be purchased. While their
comprehensiveness is a strength, this also means that they are often lengthy to complete. Broad
screeners are widely used at initial mental-health evaluation appointments or within school
system evaluations. Since scoring is often based on normative data from large samples, cutoffs
are usually available to aid interpretation of scores. These screeners can be helpful to identify
areas requiring more in-depth assessment and to quantify the magnitude of difficulty across
different areas of functioning.

One of the most common measures in this category is the Achenbach System of
Empirically Based Assessment (ASEBA) System (Achenbach, 2009), which includes caregiver
report forms (e.g., CBCL for ages 1.5-5 or 6-18 and the Adult Behavior Checklist [ABCL]), self-
report forms (Youth Self-Report [YSR] for ages 11-17 and Adult Self-Report Form [ASR]), and
SCREENING AND ASSESSMENT
a Teacher Report Form (TSR). The ASEBA system generates scores that map onto Diagnostic and Statistical Manual of Mental Disorders (DSM) criteria (DSM-oriented scales) as well as syndrome scale scores, which tap various problem behaviors, including an Aggression Scale. Although developed for the general population, the ASEBA system is widely used with individuals with ASD. A measure like the ASEBA scales that has been utilized somewhat less in autism, but has some unique strengths, is the Behavioral Assessment System for Children – 3 (BASC-3; Reynolds & Kamphaus, 2015). One disadvantage of the BASC-3 compared to the ASEBA scales is the lack of an adult version. However, advantages include its assessment of adaptive behaviors, executive functioning, and strengths in addition to coverage of the topics assessed by ASEBA. It also includes a validity index. The BASC-3 has an accompanying Behavioral and Emotional Screening System that is meant to be completed in five minutes to identify behavioral and emotional strengths and weaknesses.

Direct Screeners

One concern with informant-based measures is that challenging behavior is not actually measured in the time and place in which it is reported to occur, making it difficult to assess the specific topography of challenging behavior in question and the environmental variables responsible for its occurrence. A functional approach to screening for challenging behavior, that relies on direct observation and compliments the informant-based measures described above, was developed by Fahmie and colleagues (2016; 2020). Here, “sensitivity tests” are designed to screen for emerging challenging behavior under a set of situations that commonly occasion challenging behavior. These procedures mimic those of a functional analysis (described below) but are brief and embedded within a small-group play context. Children are exposed to brief periods of time when (1) a caregiver’s attention is diverted away from them, (2) preferred
SC histogram

**SCREENING AND ASSESSMENT**

Tangible items are restricted, and (3) demands are presented to them. If challenging behavior occurs during these situations, the caregiver either (1) delivers attention, (2) provides access to the tangible item, or (3) terminates demands. These situations are then compared with a free-play situation in which the child has access to toys and attention in the absence of demands.

Situations that occasion challenging behavior, relative to the free-play context, are then suspected to be functionally related to challenging behavior. Stated differently, the situations that produce challenging behavior likely mimic the learning environment in which challenging behavior may occur and develop. By identifying those situations, and the consequences that are likely to reinforce challenging behavior, the child can be taught alternative and more socially appropriate ways to behave and communicate under those same situations. This screening model, however, is early in its development and further research is needed to determine its efficacy.

In addition, there are also developmental assessments that rely on direct observation of behavior that more generally measure global development. For example, the Bayley Scales of Infant and Toddler Development—4th Edition (Bayley—4; Bayley & Aylward, 2019), is a standardized, norm-referenced tool that measures a child’s cognitive, language, motor, social-emotional, and adaptive development. During this assessment several tasks are given to examine how the child explores new toys, solves problems, and completes puzzles, for example. The child’s behavior is then scored and compared with other children their age to make normative comparisons of development. Although developmental assessments, such as the Bayley—4, are not specifically focused on challenging behavior, if challenging behavior is observed in a child with ASD during these assessments, pursuing assessments targeting challenging behavior should be considered.

*Standardized Measures of Challenging Behaviors*
SCREENING AND ASSESSMENT

Once a presenting concern has been identified, another class of standardized measures that focuses more specifically on challenging behavior may ensue. These scales often measure multiple forms of challenging, atypical, or disruptive behaviors, though, at times, users focus on just a single subscale related to challenging behaviors. Four of the most widely used options include the ABC-2, Emotion Dysregulation Inventory (EDI), BPI-01, and the SIB subscale of the RBS-R; key characteristics and strengths and weaknesses of these measures are described below, with some additional options included in Table 4.

The ABC-2 (Aman & Singh, 2017) is one of the most widely utilized standardized questionnaires related to challenging behaviors in ASD. It includes five broad scales (irritability, hyperactivity, lethargy, repetitive behaviors, and inappropriate speech) and has decades of support from research in autism and other intellectual and developmental disabilities for its psychometrics. The Irritability subscale, which is most commonly employed in research, includes a constellation of items tapping into tantrums, SIB, and one item on aggression. The ABC-2 Irritability subscale has been used to support FDA approval of medications to treat irritability and is widely used as an outcome measure related to challenging behaviors in ASD. The ABC-2 also includes a Hyperactivity subscale, which can be very relevant to challenging behaviors as well. The ABC-2 can be utilized across the lifespan.

The EDI (Mazefsky et al., 2018) is a standardized questionnaire of reactivity (intense negative emotional reactions) and dysphoria (low positive affect, unease). The Reactivity scale may be particularly relevant when challenging behavior is accompanied by emotional outbursts. The EDI includes a young child (2-5-year-old) and 6+ caregiver report versions; a self-report version for ages 11 and older (including adults) and adult caregiver norms are being developed. The EDI has strong evidence to support its use in ASD, as well as general community and
SCREENING AND ASSESSMENT
clinical samples, making it suitable for use in both settings specializing in ASD and those that serve broad populations including ASD. Advantages of the EDI include its validity for both nonverbal and verbal individuals, its brevity, change-sensitivity, and the availability of clinical cut-offs. The EDI is freely available for use (requests can be made at: [www.reaact.pitt.edu](http://www.reaact.pitt.edu) by completing the EDI Inquiry Form).

The BPI-01 (Rojan et al., 2001) is an informant questionnaire that measures the frequency and severity of different types of challenging behaviors. Originally developed for individuals with intellectual disability, the BPI-01 produces scores for aggressive/destructive behavior, SIB, and stereotyped behavior. It also has a short form which reduces the item number from 49 to 30 (Rojahn et al., 2012). The BPI-01 is only applicable to those who have demonstrated a behavior at least once in the past two months and is therefore more appropriate when challenging behavior is a referral concern.

The SIB scale of the RBS-R (Bodfish et al., 2000; Lam & Aman, 2007) is perhaps the most common applied standardized measure of SIB. It is part of a broader scale focused on repetitive behaviors. It is a 43-item measure that assesses behavior in the past month. It is appropriate for use with children, adolescents, and adults. Unlike the other measures, the RBS-R is freely available.
### SCREENING AND ASSESSMENT

#### Table 4

**Summary of Standardized Measures of Challenging Behavior that have been Used in or Developed for Individuals with ASD or Intellectual and Developmental Disabilities**

<table>
<thead>
<tr>
<th>MEASURE</th>
<th>Author (year)</th>
<th>Versions</th>
<th>Subscales/Diagnoses</th>
<th>Age Range</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Informant-based Screeners</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASEBA Measures (CBCL/ABCL, YSR/ASR, TRF)</td>
<td>Achenbach &amp; Rescorla</td>
<td>Parent, Self, Teacher</td>
<td>DSM-oriented (psychiatric diagnosis) scales; Syndrome scales (including Aggression); Internalizing Problems, Externalizing Problems, Total Problems</td>
<td>1.5-5 years; 6-18 years; Adult</td>
</tr>
<tr>
<td>BASC – 3</td>
<td>Reynolds &amp; Kamphaus</td>
<td>Parent, Self, Teacher</td>
<td>Child and family interaction scale, Personal adjustment and interpersonal relationships scale, School and academic performance, Self-directed problem behavior, Emotional/Internalizing and Externalizing problem behavior scales</td>
<td>2-25</td>
</tr>
<tr>
<td>(2015)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Standardized Measures on Challenging Behaviors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BPI-01</td>
<td>Rojahn et al. (2001)</td>
<td>Informant</td>
<td>Self-injury, stereotypy, aggressive/destructive behaviors</td>
<td>14-91</td>
</tr>
<tr>
<td>Children’s Scale of Hostility and Aggression:</td>
<td>Farmer &amp; Aman (2010)</td>
<td>Parent</td>
<td>Problem Scale, Provocation Scale; measures 5 domains of aggression</td>
<td>1-21 years</td>
</tr>
<tr>
<td>Reactive/Proactive</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EDI</td>
<td>Mazefsky et al. (2018, 2020)</td>
<td>Informant, Self</td>
<td>Reactivity, Dysphoria</td>
<td>Ages 2 and older</td>
</tr>
<tr>
<td>Eyberg Child Behavior Inventory</td>
<td>Eyberg &amp; Pincus (1999)</td>
<td>Parent</td>
<td>Externalizing behaviors</td>
<td>Children</td>
</tr>
<tr>
<td>Home Situations Questionnaire – Modified for ASD</td>
<td>Chowdhury et al. (2016)</td>
<td></td>
<td>Behavioral non-compliance: Socially Inflexible, Demand Specific</td>
<td>3-14 years</td>
</tr>
<tr>
<td>Nisonger Child Behavior Rating Form</td>
<td>Aman et al. (1996); Tasse et al. (1996)</td>
<td>Parent, teacher</td>
<td>Social Competence, Problem Behaviors: Conduct Problem, Insecure/Anxious, Hyperactive, Self-Injury/Stereotypic, Self-Isolated/Ritualistic, Overly Sensitive (parent)/Irritable (teacher)</td>
<td>3-16 years</td>
</tr>
</tbody>
</table>

*Note.* Achenbach System of Empirically Based Assessment (ASEBA); Adult Behavior Checklist (ABCL); Adult Self-Report Form (ASR); Behavioral Assessment System for Children – 3 (BASC-3); Child Behavior Checklist (CBCL); Emotion Dysregulation Inventory (EDI); Youth Self-Report Form (YSR); Teacher Report Form (TSR).
SCREENING AND ASSESSMENT

Psychiatric Assessments

Beyond the measures of challenging behavior reviewed above, there are psychiatric assessments that may be warranted when challenging behavior is accompanied by other symptoms of a psychiatric disorder. Individuals with ASD have high rates of co-occurring psychiatric conditions (Rosen et al., 2018). Challenging behaviors may occur in the context of, or even be due to, a wide range of psychiatric disorders. Therefore, it is important to consider assessment of psychiatric conditions, particularly in the context of a new onset of challenging behavior or a worsening of challenging behavior. Psychiatric assessment should be viewed as complementary to behavioral assessment as the identification of co-occurring psychiatric conditions should be incorporated into case conceptualization and treatment planning. There are existing resources that cover the topic of psychiatric conditions in ASD in depth (i.e., The Oxford Handbook of Autism and Co-Occurring Psychiatric Conditions, White et al., 2020). Below we cover some of the key themes regarding psychiatric assessment in ASD.

Notably, there are no preferred or gold-standard validated mental-health assessments that reliably identify co-occurring psychiatric conditions in addition to ASD. In general, structured interviews are considered the most comprehensive psychiatric diagnostic assessments, but they can be time consuming. The Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime version (Kaufman et al., 1997) and Structured Clinical Interview for DSM-5 (First et al., 2014) are perhaps the mostly widely used structured psychiatric interviews in general; both can take several hours to complete and require a substantial amount of training in administration and score. The MINI International Neuropsychiatric Interview (Lecrubier et al., 1997) has been considered in ASD with promising psychometrics (Mosner et al., 2019). Although additional evaluation is needed, it may offer a brief (~15 min) alternative to
SCREENING AND ASSESSMENT
longer structured interviews. While some interviews have been developed for intellectual and
developmental disabilities and ASD populations (e.g., Autism Comorbidity Interview [Leyfer et
al., 2006]; Baby and Infant Screen for Children with Autism Traits [Matson et al., 2007]), they
are also not in widespread use and are more likely to be utilized in research contexts.

Questionnaires may be a more feasible initial screen for psychiatric diagnoses. The Child
and Adolescent Symptom Inventory (Gadow, 2015) has been widely utilized in ASD and is
designed to map onto DSM-5 criteria. It is a standardized behavior rating scale for children ages
5 to 18. It provides symptom count cutoff scores, symptom severity scores, impairment cutoff
scores, and clinical cutoff scores for all major DSM-5 disorders. The previously described
ASEBA measures and BASC-3 also provide indices of psychiatric diagnoses and can be useful
as psychiatric screening questionnaires. One caution is that it is not uncommon for individuals
with ASD to score high on the CBCL and BASC-3 psychiatric diagnoses indicators even when
not meeting criteria for co-occurring diagnoses. It is possible that these measures also capture
common impairments seen in ASD even when a separate diagnosis is not warranted. This may
reflect difficulty in the differential diagnosis of psychiatric symptoms in ASD as well as
symptom overlap (e.g., symptoms of withdrawal in both ASD and depression, repetitive actions
in both ASD and OCD, etc.).

Further, the Academic Autism Spectrum Partnership in Research and Education
(AASPIRE), a partnership between researchers and autistic adults, is developing a toolbox of
mental health measures for adults, many of which have been modified to be more appropriate
and valid in ASD (Nicolaidis et al., 2020). There are also some recent and ongoing initiatives to
develop new psychiatric symptoms measures for autism (e.g., the Pediatric Anxiety Rating
SCREENING AND ASSESSMENT
Scale; Maddox et al., 2020). As such, the assessment of mental health concerns in ASD is an area with anticipated future growth.

Summary

Standardized, informant-based questionnaires can be useful for efficiently determining the presence and degree of challenging behaviors. As can be seen in review of Tables 1-3, these measures are underutilized. Often, studies employ single items or custom measures that lack reliability, validity, and sensitivity to change. The measures described above and summarized in Table 4 offer a starting point for the standardized assessment of challenging behaviors in ASD. The specific measures employed in practice should be based on clinical expertise, empirical evidence, and tailored to the individual’s needs and preferences (APA, 2006). Nonetheless, most options produce summary scores for clusters of challenging behaviors (i.e., an externalizing behaviors score, irritability score) versus specific scores for different types of challenging behaviors (e.g., aggression, SIB). Of note, we were unable to identify any standardized measure of elopement.

Future research should focus on the development of efficient, precise, and change sensitive standardized assessments of specific challenging behaviors across the lifespan as well as consideration of the psychometrics of standardized questionnaires that are utilized in non-ASD populations. For example, the Multidimensional Assessment of Preschool-Disruptive Behavior Scale (MAP-DB; Wakschlag et al., 2014) is a particularly psychometrically strong measure of challenging behaviors (Mazefsky et al., 2021). It offers versions now available across the childhood years (Infant/Toddler, Preschool, School Age, and Adolescence; Biedzio & Wakschlag, 2018), including early childhood which is a gap in measure options of specific challenging behaviors in ASD and is a particularly difficult period of assessment given the
SCREENING AND ASSESSMENT
occurrence of dysregulation even among normative samples at that age. The MAP-DB is widely utilized outside of ASD but has thus far not been considered within the ASD community.

Two other directions for measurement include the use of self-report measures and a focus on non-suicidal self-injury. All the measures listed in Table 4 require informant report. However, many adolescents and adults with ASD have the verbal abilities to complete measures themselves. For instance, the Buss-Perry Aggression Questionnaire (Buss & Perry, 1992; Buss & Warren, 2000) is a self-report questionnaire of aggression, verbal aggression, hostility, and anger that has been used in hundreds of studies outside of ASD. Understanding its utility in ASD is worth consideration. There is also recent interest in considering measures of non-suicidal self-injury in ASD, which may offer important insights into SIB in ASD, particularly among more cognitively able individuals.

When challenging behavior is associated with comorbid psychiatric symptoms, psychiatric assessments should also be considered to complement behavioral assessments. Comorbid psychiatric conditions are common in ASD and should be integrated into case conceptualization and treatment planning related to challenging behaviors. It is worth noting that considerable resources also exist for the assessment of prosocial behaviors. Understanding individual strengths and specific behavioral repertoires that may lessen or even prevent the development of challenging behaviors is critical to the ongoing comprehensive developmental care for individuals with ASD. For instance, the Values in Action (VIA) Inventory of Strengths has shown promise in terms of identifying prosocial aspects of character among persons with intellectual and developmental disabilities (Shogren et al., 2018). Once identified, well-being may be promoted by leveraging the individual’s strengths rather than focusing on reducing deficits. Although merging the disability field with positive psychology is quite
SCREENING AND ASSESSMENT
nascent, it does hold promise for improving academic, social, and mental-health outcomes (Lai et al., 2018; Vuorinen et al., 2018; Raley et al., 2020) and should be considered in future research.

Finally, direct screening procedures of challenging behavior rely on observation of behavior during situations that commonly contribute to the occurrence and maintenance of challenging behavior in the ASD population. This structured observational approach to screening holds much promise in those situations for which children are sensitive to reactions (i.e., engage in emerging forms of challenging behavior) and can be immediately targeted for prevention-level treatment. For example, a child may be taught to engage in context-specific communication that would deter further development of challenging behavior occasioned by those situations. The long-term benefits of this approach are unknown at this time but warrant future research. Nevertheless, the notion of assessing the environmental variables that may contribute to the development of challenging behavior highlights the goal of function-based assessment as well, which is described next.

Function-Based Assessment

Once challenging behaviors have been identified, timely and accurate assessment of the level, severity, and maintaining variables related to the challenging behavior is critical to matching appropriate type and levels of care. Of key importance is identifying the target behavior in need of treatment and the environmental variables responsible for its maintenance. Standardized assessments, like those described above, are helpful in making normative comparisons; however, they are limited in their ability to detect the environmental variables that contribute to challenging behavior. As a complement to these standardized approaches, indirect, direct, and experimental assessment procedures have been developed and packaged into what has been termed functional behavior assessment (FBA). These assessment methods provide
SCREENING AND ASSESSMENT
a means for objective measurement of challenging behavior, understanding of behavioral function, and a prescription for later development of intervention.

FBAs are comprised of various procedures that assess the environmental variables suspected to evoke and maintain challenging behavior in the natural environment. These assessment procedures are based on a number of key assumptions (Dunlap et al., 1991; Horner & Carr, 1997; Martens & Ardoin, 2010; O’Neill et al., 1997; Sullivan et al., 2021): (a) the focus of the assessment is on challenging behavior itself, rather than viewing challenging behavior as a sign indicating an underlying disorder, (b) challenging behavior varies systematically across environmental situations and has been learned from past experiences, (c) through repeated measurement predictable patterns of challenging behavior can be identified, and (d) the contingencies supporting challenging behavior that are identified through functional assessment can be modified during treatment. Overall, the purpose of an FBA is to improve effectiveness and efficiency of the behavioral treatment (Horner, 1994), and treatments based on the function of challenging behavior have consistently been shown to be more effective than non-function-based interventions (Didden et al., 1997; Iwata et al., 1994; Saini et al., 2021).

The assessment procedures available to practitioners when conducting an FBA are not static and prescribed. Instead, they are intended to be a menu of techniques available to practitioners to arrive at the identification of the specific controlling variables that maintain a behavior, and to use that information to guide selection of evidence-based interventions (see Evidence Based Practices document). For example, if physical aggression is found to be maintained by obtaining access to preferred activities, then in treatment the individual may be taught to request preferred activities using a more socially appropriate, communicative response (Carr & Durand, 1985; Saini & Sullivan, 2021). It should also be noted that FBAs are helpful in
SCREENING AND ASSESSMENT
identifying individual strengths and preferences that may be incorporated in treatment. For instance, preference assessments (e.g., Fisher et al., 1992; DeLeon et al., 1996; Roane et al., 1998) may be used to identify reinforcers specific to the individual, and through direct observation adaptive skills may be highlighted. For present purposes, however, the most common strategies used to assess the function of challenging behavior will be reviewed and have generally fallen into three categories (Roane et al., 2019): (1) indirect assessment, (2) descriptive assessment, and (3) functional analysis. In the following sections we will briefly describe common FBA strategies to clarify how functions of challenging behavior are assessed.

**Indirect Assessment**

Indirect functional assessment (Gadaire et al., 2021) describes a group of procedures that aim to efficiently gather information about an individual’s challenging behavior and the events that surround its occurrence. More specifically, these procedures focus on identifying the antecedent events that precede challenging behavior and the responses that follow challenging behavior which may serve as reinforcement. Common procedures include record reviews (e.g., medical, school), behavioral interviews (structured or semi structured), and various checklists, questionnaires, and rating scales. Below, commonly used indirect assessments to hypothesize function(s) of challenging behavior (Gadaire et al., 2021) are reviewed.

The Questions About Behavioral Function (QABF; Matson & Vollmer, 1995) is a rating scale designed to assess possible functions of challenging behavior. The QABF contains 25 items that correspond with five potential sources of reinforcement: attention, escape, non-social (automatic-positive), physical (automatic-negative), and tangible. Items are scored based on how often the challenging behavior is reported to occur across situations using a 4-point Likert-type scale (0=Never, 3=Often). Matson et al. (2012) conducted a review of the QABF and reported
SCREENING AND ASSESSMENT

good test-retest reliability, acceptable inter-rater reliability, and good internal consistency. Additionally, conclusions drawn from the QABF were found to be like those from functional analyses.

The Functional Analysis Screening Tool (FAST; Iwata & DeLeao, 1996) is a self-reported 16-item questionnaire designed to identify antecedent and consequent events that may contribute to the occurrence and maintenance of challenging behavior. The 16 items are categorized into four sections that describe the conditions under which the behavior occurs. Relevant stakeholders (i.e., caregivers, teachers) complete a structured questionnaire designed to gather information about functional characteristics of the problem behavior to analyze potential sources of reinforcement. Although Iwata et al. (2013) found the FAST to produce somewhat reliable and valid reports, it is not recommended that practitioners conduct the FAST without additional descriptive and/or experimental assessments to inform function.

The Motivation Assessment Scale (MAS; Durand & Crimmins, 1988), is a 16-item checklist used to develop hypotheses regarding behavioral function. Informants rate how often the individual engages in challenging behavior using a 7-point Likert-type scale ranging from 0 (Never) to 6 (Always). Item ratings are summed and categorized by potential sources of reinforcement: sensory, escape, attention, and tangible. Durand and Crimmins reported evidence of good test-retest reliability over a 30-day period ($r = .89-.98$), adequate interrater reliability ($r = .66-.92$), and good predictive validity between ranked scores on the MAS and functional analysis outcomes. However, others have reported suboptimal psychometrics (see Sigafoos et al., 1994; Zarcone et al., 1991).

O’Neill et al. (1997), developed an indirect-assessment system titled, The Functional Analysis Interview (FAI). The FAI is a structured interview that includes eleven sections
SCREENING AND ASSESSMENT
designed to identify potential functions of challenging behavior. The FAI takes approximately
45-90 minutes to complete and provides a guide for the interviewer across content areas. For
example, the FAI helps to define the challenging behavior, identify the antecedents and
consequences for challenging behavior, determine the individual’s communicative abilities,
identify potential reinforcers, and review the history of previous interventions.

Each of these indirect assessments allow for hypotheses regarding the function(s) of
challenging behavior to be developed based on the environmental events that were reported to
occasion challenging behavior. These procedures are practical and efficient, however, because
the outcomes are based on the informant’s perception and recall of past events, potential biases
in reporting may lead to inaccurate identification of function (Gadaire et al., 2021; Iwata et al.,
2013). To address these concerns, direct observation and measurement of challenging behavior is
needed. Below, various behavioral measurement strategies are described with their relative
strengths and weaknesses.

Direct Measurement

This section outlines strategies for directly measuring and recording challenging
behavior. Each strategy has its own strengths and limitations. We do not provide direct
recommendations that one procedure should be used over another. Instead, we highlight each
of the strategies benefits and drawbacks, and provide readers with a decision tree for selecting an
appropriate measurement system. Key to any direct and reliable measurement of challenging
behavior is the development of operational definitions for the target responses. Definitions must
be objective and clear, with well-established understanding of when a response begins and ends.
The operational definition must also include examples and non-example such that any observer
SCREENING AND ASSESSMENT
clearly understands what a behavior is and is not. Table 5 provides a list of common
topographies and corresponding operational definitions for challenging behaviors.

Table 5

Sample Operational Definitions of Common Challenging Behaviors

<table>
<thead>
<tr>
<th>Response Topography</th>
<th>Operational Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical aggression</td>
<td>Any completed, attempted, or blocked response that could cause injury to another person. This includes but is not limited to slapping, scratching, kicking, pinching, pushing, head butting, and throwing objects at people. Examples include: forceful contact of hand (open or closed) or arm (with or without another object) against any part of the therapist’s body. Non-examples include: giving high-five, giving someone a hug.</td>
</tr>
<tr>
<td>Self-injury</td>
<td>Any completed or blocked response that is self-directed such that repetition of the behavior over time has or will cause bodily injury. This includes but is not limited to head banging, self-hitting, biting, eye-poking, hair pulling and pinching. Examples include: forceful contact or attempted contact of client’s hand (open or closed, with or without object) or foot against any part of the client’s own body from at least 2 inches away or greater. Non-examples include: scratching head, tapping foot against floor.</td>
</tr>
<tr>
<td>Property destruction</td>
<td>Any completed or blocked response that could cause damage to materials or any other objects or surfaces within the immediate environment. This includes throwing objects, kicking/hitting objects, over-turning furniture, climbing on objects, and swiping objects from a table or other surface. Examples include: Patient projects an object from a distance of at least 6 inches with force (not directed at therapist). Non-examples include: Playing catch during play or bumping into table and knocking off materials.</td>
</tr>
<tr>
<td>Elopement</td>
<td>Any completed, attempted, or blocked instance of a patient leaving a supervised. Examples include: moving from a supervised room or area without permission or moving more than 5 feet away from the therapist. Non-example includes leaving assigned area when instructed to do so.</td>
</tr>
<tr>
<td>Flopping</td>
<td>Any completed, attempted, or blocked instance in which the client’s body falls from a standing position to the floor or ground such that his or her midsection (i.e., back, buttocks, stomach, or shins) contacts the floor or ground). Non-example includes laying on floor playing with toy or watching television.</td>
</tr>
</tbody>
</table>

Once operational definitions are developed, strategies for direct measurement should be selected. There are many measurement strategies available to practitioners each with strengths and limitations. Table 6 provides a summary of the measurement procedures discussed below with associated strengths, potential limitations, and examples of use from the published literature. These data may be collected through a variety of means including paper and pencil data collection, use of response clickers and timers, or computer-based data collection software.
SCREENING AND ASSESSMENT
(e.g., BDataPro; Bullock, 2017). Deciding what data-collection strategy to use and the manner in which challenging behavior should be measured are important considerations. Thus, a decision-making tree for determining the appropriate measurement strategies is provided in Figure 4.

Continuous measurement. Continuous measurement requires constant observation and, therefore, may not be possible to conduct in applied settings. Although labor-intensive, continuous measurement systems provide a comprehensive and ongoing account of the behavior of interest. Frequency, duration, latency, and intensity are the most commonly measured dimensions of behavior in which continuous measurement procedures are used.

Frequency measures attempt to capture the repeatability dimension of a target response and require minimal instrumentation. To conduct a frequency measure, one simply counts the occurrence of the target response. One could also divide the total count of the target response by the duration of the observation period to produce response rate. Response rate is an important frequency summary measure because it controls for unequal observation periods. That is, in scenarios where session durations are not constant, comparing frequency without considering observation time may skew the data; response rate equates across these different observation durations.

Duration recording is used to capture the temporal extent, or the time that passes between the onset and offset of a target response. To use a duration measure, one must identify the onset and offset criteria, then begin a timing device (e.g., stopwatch or timer) when onset criteria are met and stop the timing device when offset criteria are met. Duration is helpful when targeting responses with a long temporal extent. It may also be beneficial for responses with a brief temporal extent, which also have a rapid rate of occurrence. Duration can be summarized in several ways. The measure can be summarized with the frequency of the target response to
SCREENING AND ASSESSMENT
produce the duration per occurrence of target responses. The total duration of a target response can be divided by the observation period to derive a percentage duration. Lastly, the duration per occurrence of the target response can be averaged to determine the mean duration of the response.

Latency is similar to duration but captures the temporal locus of two events. Latency is typically used to identify when a target response occurred in relation to some other environmental event. For instance, if an instruction is given to a student, and then she aggresses toward the instructor, the time that elapses between instruction and challenging behavior is the response latency. To conduct this measurement procedure, clear criteria are needed for when the observer is to start and stop their timing device. Interresponse time is a variation of the latency procedure where, instead of timing the latency from evocative stimulus (e.g., instruction) to target response (e.g., challenging behavior), one records the latency between the cessation of one target response to the onset of the next.

Intensity recording captures the magnitude of a target response. This measurement strategy may be more difficult to conduct in applied settings due to instrumentation requirements. Intensity is measured depending on the topography of the target response. For instance, if an individual engages in screaming behavior, intensity may be captured by decibel meter. If the target response includes challenging behavior toward others, intensity could be measured by a pressure plate (though, to date, intensity measures of challenging behavior have been understudied). It is often easier to measure intensity by the permanent products the target response leaves after its occurrence or by rating scale. Figure 1 provides an example of a rating scale that may be used to gather information regarding intensity or severity of challenging behavior.
SCREENING AND ASSESSMENT

Figure 1

Sample Rating Scale Depicting Varying Levels of Challenging-Behavior Severity

Severity Range (check all that apply)

<table>
<thead>
<tr>
<th>Level 1</th>
<th>Challenging behavior resulting in (a) no marks on body and (b) no blows close to or contacting the eyes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 2</td>
<td>Challenging behavior resulting in (a) reddening of skin, and/or (b) mild swelling</td>
</tr>
<tr>
<td>Level 3</td>
<td>Challenging behavior resulting in (a) light scratches, (b) small or shallow breaks in skin, and/or (c) moderate to severe swelling</td>
</tr>
<tr>
<td>Level 4</td>
<td>Challenging behavior involving blows close to or contacting the eyes or resulting in (a) scratches that leave scars, (b) breaks in skin that leave scars, and/or (c) trauma resulting in broken bones or lasting tissue damage or disfigurement</td>
</tr>
</tbody>
</table>

This rating scale is a less desirable measure to determine the occurrence of challenging behavior as observers must make a subjective evaluation of the intensity of the target response. Thus, it is not suggested that rating scales be used alone in measuring challenging behavior unless there is no other option. In resource-limited situations, it may be best to use a visible product of the response (called permanent-product recording), which would allow practitioners to gather additional information related to the frequency and intensity of target challenging behaviors. It is important to note that this permanent-product data is not a direct measure of the target response and requires a degree of inference, which sacrifices some accuracy and validity of these data. The tactics of measuring permanent products vary depending on the topography of interest. For example, when collecting the pictorial examples of a sustained bite mark provided in Figure 2 permanent-product data would include the outcome (i.e., tissue damage) that resulted from the bite mark; however, additional information of the aggressive act, such as the number of bites (frequency) that occurred would not be represented.
Other adaptations of established measures for recording damage produced by a response may also be relied upon. For instance, SIB may be measured via description of surface tissue damage via the Self-injurious Trauma (SIT) Scale (Iwata et al., 1990). Although the SIT Scale may be used to document the surface tissue damage targeted toward implementers of behavioral interventions, that is not its original intention.

**Discontinuous measurement.** Discontinuous measures are most frequently employed when a practitioner desires to record sample measures of a target response during a prescribed observation period. Discontinuous measurement procedures require dividing an observation period into equal intervals and recording the occurrence of responses across those intervals. Discontinuous measures are most appropriate for target responses where the dimension of interest is repeatability and temporal extent. Relative to continuous measurement that requires constant observation, discontinuous measurement breaks down the observation into specific intervals and may not capture every instance of behavior thereby producing only an estimate of the occurrence of the target response.
SCREENING AND ASSESSMENT

The primary benefit of discontinuous recording procedures is that they are typically easier to use than frequency or duration measures as they do not require precision in the recording of the exact occurrence of the target response. Three main types of discontinuous measurement are common: (1) partial-interval recording, (2) whole-interval recording, and (3) momentary time sampling. When summarizing the following discontinuous measures, it is standard practice to report the occurrence of the target response in percentage of intervals or percentage of the observation period.

Partial-interval recording involves recording whether the target response is observed or is not observed during a specified time interval. After an observation time is identified, the timeframe is divided into smaller intervals of equal length. Partial-interval recording is likely suitable to record high-rate behavior across multiple forms (LeBlanc, 2016). Partial-interval recording often overestimates the occurrence of a target response.

Whole-interval recording involves first dividing the observation period into equal intervals. The occurrence of the target response is only reported if the target response persists for the entirety of an interval. Due to the requirement that the response persists for the whole interval, it tends to underestimate the occurrence of the target response. Because of this underestimation, it is not advised that whole-interval recording be used for responses which are targeted for reduction. To illustrate whole-interval recording, suppose we again divided a one-minute observation window into 6, 10-second intervals, we might observe the student to engage in challenging behavior at least once in 4 of the 6 intervals but only throughout the entire interval twice. Given this, whole-interval recording of her challenging behavior would be 33.3%.
SCREENING AND ASSESSMENT

The final common discontinuous measurement procedure is momentary time sampling. Like the above discontinuous procedures, the observation period is split into equal observation intervals. For most of the interval, the observer does not look for the occurrence of the target response, but rather looks up at the end of each interval for an observation check of approximately 1-3 s. If the target response occurs at any point during the observation check, the target response is recorded for the entire interval. Momentary time sampling has the potential to both over- and underestimate the occurrence of target responses. These errors are typically a function of the observation interval length and characteristics of the target response. Fiske and Delmolino (2012) provide a more in-depth discussion of the factors that impact error rate in momentary time sampling than what is possible in this document. Suppose that student’s aggressive behavior is being observed in her classroom. She might be observed for 3 seconds at the end of each 10-second interval but not during the remaining 7 seconds of each interval. As with the previous examples, if the student engaged in some challenging behavior during 4 of 6 total intervals, but only three of those occurrences in the 3-second observation window, the momentary time sampling measure would reveal challenging behavior to occur during 50% of the observation (i.e., underestimation relative to partial-interval and overestimation relative to whole-interval). Table 6 provides a summary of the aforementioned measurement procedures with associated strengths, potential limitations, and examples of use from the published literature. Figure 3 provides a decision-making model for selecting appropriate measurement procedures.
<table>
<thead>
<tr>
<th>Measurement procedure</th>
<th>Description</th>
<th>Resulting measure(s)</th>
<th>Strength(s)</th>
<th>Potential limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Event recording</td>
<td>Record each instance of behavior as it occurs</td>
<td>Frequency, rate, percentage of opportunities (for restricted operants)</td>
<td>Direct measures of the behavior</td>
<td>Requires constant vigilance; may be impractical for high-frequency or non-discrete behavior</td>
</tr>
<tr>
<td>Duration recording</td>
<td>Record the amount of time from onset to offset for each behavior as it occurs</td>
<td>Total duration, mean duration, percentage duration</td>
<td>Direct measures of the behavior (including frequency)</td>
<td>Requires constant vigilance; requires a timing device</td>
</tr>
<tr>
<td>Latency recording</td>
<td>Record the amount of time that passes between the discriminative stimulus and the onset of the behavior</td>
<td>Mean latency</td>
<td>Direct measures of the temporal relation between the discriminative stimulus and the behavior</td>
<td>Requires constant vigilance; requires a timing device</td>
</tr>
<tr>
<td>Intensity recording</td>
<td>Record a dimension of intensity (e.g., force, volume) for each instance of the behavior</td>
<td>Various (e.g., mean decibel, mean rating per event)</td>
<td>Direct measures of the behavior; automated recording possible with some dimensions (e.g., volume)</td>
<td>Requires a reliable and valid measurement device (e.g., decibel meter) or rating system; requires constant vigilance</td>
</tr>
<tr>
<td>Permanent-product recording</td>
<td>Document the effects of a behavior on the environment</td>
<td>Various (e.g., wound size, number of holes in wall)</td>
<td>Allows measurement of behavior that occurs at inaccessible times</td>
<td>An indirect assessment of behavior; behavior must reliably produce the product and be the only source of the product</td>
</tr>
<tr>
<td>Partial-interval recording</td>
<td>Record whether a behavior occurred at all during specific time intervals for defined observation period(s)</td>
<td>Percentage of intervals in which the behavior occurred</td>
<td>Does not require constant vigilance</td>
<td>Generates an estimate of behavior; generates systematically overestimates the occurrence of behavior; requires a timing device</td>
</tr>
<tr>
<td>Momentary time sampling</td>
<td>Record whether a behavior occurred at a given moment for defined observation period(s)</td>
<td>Percentage of samples or intervals in which the behavior occurred</td>
<td>Allows concurrent measurement of multiple individuals or behaviors; does not require constant vigilance; good correspondence to event recording compared to other discontinuous procedures</td>
<td>Generates an estimate of behavior. Inappropriate for short-duration or low-frequency behavior; requires a timing device</td>
</tr>
</tbody>
</table>
SCREENING AND ASSESSMENT

Figure 3

A Decision-Making Model for Selecting Data Measures

Note. Adapted by W. E. Sullivan from LeBlanc et al. (2016). A proposed model for selecting measurement procedures for the assessment and treatment of challenging behavior.
SCREENING AND ASSESSMENT

Descriptive Assessment

Once challenging behavior is operationally defined and a measurement system is selected, descriptive assessments (Castillo et al., 2018; Lerman & Iwata, 1993; Mace & Lalli, 1991; Martens et al., 2008) may ensue. Descriptive assessments utilize direct observation and data collection of the challenging behavior under naturalistic environmental conditions, thereby addressing the concerns with informant reports of past events. One assessment tool, the scatterplot (Touchette et al., 1985), can be used to visually depict the occurrence and temporal relationship of the response(s) to various environmental events illuminating possible patterns not readily observed in a graph or extrapolated from indirect assessments. Additionally, descriptive assessments can be helpful in generating hypotheses regarding the function of challenging behavior (Lerman & Iwata, 1993). Descriptive assessments provide information about the environmental situations under which challenging behavior does and does not occur, and most importantly, provides crucial information needed to develop socially valid experimental analyses of the putative variables influencing challenging behavior (i.e., functional analysis, described below).

Within the context of FBAs, descriptive assessment procedures typically begin by observing problem behavior across different antecedent conditions (Erchul & Martens, 2010) referred to as scatterplot recording (Touchette et al., 1985). Scatterplot recording examines under what conditions challenging behavior is most likely to occur but does not measure the consequences that follow challenging behavior. Thus, hypotheses regarding behavioral function based on scatterplot recordings are limited. Scatterplots may still be informative, however, in
SCREENING AND ASSESSMENT
determining the optimal time to engage in recording of behavior and its consequences (Eckert et
al., 2005).

Another way to determine under what conditions challenging behavior is most likely to
occur while simultaneously examining consequences is to engage in Antecedent-Behavior-
Consequence (A-B-C) recording (Bijou et al., 1968). This type of assessment involves recording
the occurrence of problem behavior, under what conditions it occurred (antecedents), and what
consequence(s) were provided. This process continues until a clear pattern of antecedents and
consequences associated with problem behavior emerges (Lee & Miltenberger, 1997).

A-B-C recording is advantageous in that it can provide descriptive information in a
systematic manner about the events that surround behavior. For example, Tustin (1995) utilized
A-B-C recording procedures to determine possible functions of stereotypy in a 28-year-old male
diagnosed with autism. Results suggested that stereotypy was associated with changes between
work activities (e.g., packing materials). However, there are several limitations with A-B-C
recordings outlined by Iwata et al. (2000). First, because A-B-C recordings typically do not
provide operational definitions for each antecedent and consequence, their reliability is
questionable. Second, there is no uniform way to summarize and interpret the data, which may
produce subjective and biased conclusions. Finally, because data collection only focuses on
problem behavior, frequently delivered consequences (e.g., attention) may follow problem
behavior by chance, leading to an inaccurate functional hypothesis (Thompson & Iwata, 2007).

An alternative strategy for examining the relationship between behavior and its
consequences is to conduct sequential recordings and examine the conditional probability of a
consequence given behavior. This type of assessment typically involves recording behavior and
its consequences in brief (e.g., 10 s) intervals as they occur in sequence throughout an
SCREENING AND ASSESSMENT
observation period (Vollmer et al., 2001; Martens et al., 2008). Prior to collecting these data, specific challenging behavior(s) and consequences are defined so that behavior categories (i.e., challenging behavior and all other behavior) are mutually exclusive and consequences represent broad categories of reinforcement (i.e., social-positive, social-negative, automatic positive or negative). Following data collection, conditional probabilities are calculated and those consequences that have a high probability of following challenging behavior indicate potential functions.

Functional Analysis

Although indirect and descriptive FBA strategies are helpful in identifying patterns of challenging behavior, these procedures fall short in being able to demonstrate a functional relation between challenging behavior and the environmental events suspected to produce it. That is, although hypotheses regarding the function of challenging behavior can be formed these assessments do not allow for those hypotheses to be confirmed. To address this issue, Iwata and colleagues (1982/1994) developed a functional analysis that systematically manipulated the environmental variables hypothesized to occasion self-injurious behavior within a single-case experimental design. In a functional analysis, various test conditions are designed to test specific hypotheses regarding behavioral function, which are then compared with a control condition to demonstrate a functional relation. In each test condition, a specific situation that is suspected to evoke challenging behavior is presented and contingent on challenging behavior a specific consequence is provided that may increase the chances of challenging behavior occurring again in the future under similar situations (i.e., reinforcement). The relative benefit of functional analysis over other functional assessment procedures discussed above is that the clinician has direct control over the contingencies that are influencing challenging behavior, which permits a
SCREENING AND ASSESSMENT
more detailed level of analysis and hypothesis testing (Vollmer et al., 2012). Below, common
test and control conditions (see Saini et al., 2021) are described in Table 7.

Table 7

Description of Functional Analysis Test and Control Conditions

<table>
<thead>
<tr>
<th>Test Conditions</th>
<th>Antecedent Situation</th>
<th>Consequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attention</td>
<td>Diverted or divided attention (e.g., reading a magazine or talking to another person).</td>
<td>Social disapproval or reprimands (e.g., “Do not do that. It is not nice.”).</td>
</tr>
<tr>
<td>Tangible</td>
<td>Restricted access to preferred activity or toys (e.g., removal of preferred toy).</td>
<td>Access to restricted activity or toy.</td>
</tr>
<tr>
<td>Escape</td>
<td>Presentation of nonpreferred or aversive situation (e.g., presentation of academic demands or loud noises).</td>
<td>Removal of the aversive situation (e.g., a break from work or removal of noise).</td>
</tr>
<tr>
<td>Ignore/ Alone</td>
<td>No interaction.</td>
<td>Sensory outcomes of the challenging behavior.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Control Condition</th>
<th>Antecedent Situation</th>
<th>Consequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Toy play</td>
<td>Near continuous attention and access to preferred tangible item in the absence of aversive stimulation.</td>
<td>No differential consequences are provided.</td>
</tr>
</tbody>
</table>

Since the initial development of the functional analysis procedure by Iwata et al. (1982/1994), its use has been validated across hundreds of studies targeting various topographies and severity levels of challenging behavior (Beavers et al., 2013; Hanley et al., 2003). Furthermore, it has been modified in a number of ways to increase efficiency (Falcomata et al., 2016), to develop novel conditions (Owen et al., 2020; McCord et al., 2001), and to implement across a variety of settings (e.g., school; Bloom et al., 2013). Nonetheless, functional analyses of challenging behavior require specialized training and inclusion of safety precautions to minimize
SCREENING AND ASSESSMENT
risk (e.g., session termination criteria, personal protective equipment; Weeden et al., 2010; Saini et al., 2021) and may not be appropriate in all situations. Ultimately, by identifying the function(s) of challenging behavior, the contingencies of reinforcement that contribute to the occurrence of challenging behavior can be directly altered in treatment. Thus, identifying the function of behavior is directly prescriptive of treatment.

Summary

The goal of FBAs is to identify the antecedent conditions that evoke challenging behavior and the consequences that maintain it. By identifying the reinforcement contingencies that occasion challenging behavior, treatment can be arranged to directly affect those contingencies and reduce challenging behavior. Utilization of informant report (i.e., indirect assessment) and measurement of challenging behavior via direct observation (i.e., descriptive assessment) can lead to the development of functional hypotheses. Only functional analysis, however, can confirm those hypotheses by manipulating the antecedent and consequent variables within an experimental arrangement.

The specific FBA methods (e.g., indirect assessments, functional analysis) employed will ultimately depend on the practice setting and training of the individual conducting the assessment. These decisions should be rooted in evidence-based practice that includes consideration of empirical evidence, clinical expertise, and the client’s preferences and values (APA, 2006). Although functional analysis can confirm the function(s) of challenging behavior, for example, it may not be appropriate for every case and in all settings (e.g., dangerous forms of challenging behavior; limited resources or training). For example, if a young child with ASD begins to engage in disruptive behavior in the classroom, a school-based FBA that utilizes indirect assessments to develop functional hypotheses and design intervention may be sufficient.
SCREENING AND ASSESSMENT
However, if treatment is ineffective and challenging behavior persists or worsens, a more sophisticated functional analysis may be needed to confirm which contingencies are functionally related to the child’s challenging behavior. In the following sections, we will outline a framework for identifying and assessing challenging behavior that can be used to inform the level of care needed over time.

Multi-Tiered Systems of Support

As noted above, individuals with ASD that engage in challenging behavior and their families may interact with multiple support systems from the time a child is born through adulthood. To date, however, there are no uniform screening measures for youth with ASD in relation to development of challenging behavior or a framework for connecting them with appropriate treatment services and monitoring their progress across the lifespan. Take for example a child with ASD that engages in their first severe tantrum. The child’s caregiver may reach out to their pediatrician and school system for help. Perhaps this tantrum was an isolated incident, and no further intervention is necessary. On the other hand, perhaps this tantrum set the stage for an escalating pattern of challenging behavior across school and community settings. In either case, it is imperative that the child’s behavior be monitored to either confirm that no further intervention is needed or to provide appropriate referrals to specialized treatment providers. For this reason, a coordinated system of care for (a) screening for challenging behavior in individuals with ASD, (b) connecting them with appropriate treatment services, and (c) surveillance of these individuals over time is needed.

An example of such a framework is termed a multi-tiered system of support (MTSS). An MTSS is an evidence-based framework that takes a data-based problem-solving approach to integrate assessment and intervention across a continuum of services (Sugai & Horner, 2009).
SCREENING AND ASSESSMENT
The MTSS described here is based on the decades of work that have been done on response-to-intervention, a commonly adopted model in schools to promote positive behavior change (e.g., Caplan, 1964; Sugai et al., 2000; Sugai & Horner, 2009; 2010; Walker et al., 1996), and mimics the triage systems in many medical interventions (e.g., seek care when problem emerges, begin with general pediatrician, and refer to specialist). An MTSS is designed to combat the “wait it out” approach by which caregivers are advised to wait to see if skills develop or symptoms subside. This “wait it out” approach is chronically characterized by delayed service delivery. That is, rather than waiting for severe forms of challenging behavior to emerge, and then obtaining treatment services, MTSS utilizes on-going screening and assessment measures to more immediately identify individuals in need of intervention and links assessment outcomes to evidence-based treatments (see Evidence Based Practices document).

Thus, the primary purpose of a MTSS is to prevent the development of more severe forms of challenging behavior in the ASD population. The National Research Council and Institute of Medicine conducted an extensive review on interdisciplinary research on prevention and suggested that childhood behavior disorders are “preventable” (O’Connell et al., 2009, pp. xii-xiv; Fahmie et al., 2020) suggesting that early identification and targeted treatments can mitigate the occurrence and development of challenging behavior. Since challenging behavior may persist over time, prevention should be of high priority to avoid costly and harmful outcomes (Waddell et al., 2018). A MTSS addresses prevention from a problem-solving approach (e.g., Lewis et al., 2010) by providing a continuum of supports that utilizes empirical data to select, evaluate, and monitor the effects of interventions.

For example, if a 3-year-old child with ASD is reported to engage in mild forms of challenging behavior (severity levels operationalized below), a MTSS would provide a
SCREENING AND ASSESSMENT
framework for (a) detecting the presence of challenging behavior using empirically validated screening methods, (b) linking the outcomes to evidence-based early intervention programs (Peters-Scheffer et al., 2011) to prevent the further development of challenging behavior, and (c) monitoring progress over time with on-going modifications to treatment as needed. Thus, if treatment is not found to produce socially significant reductions in challenging behavior, an MTSS can facilitate the direction of increasingly individualized and specialized intervention.

Primary-Care Pathway

As an example of how an MTSS system may be utilized to address challenging behavior in the ASD population, McGuire et al. (2016) developed a practice pathway designed to help primary care providers (PCPs) screen, assess, coordinate treatment, and monitor challenging behavior displayed by their patients with ASD in collaboration with parents, schools, and specialized-care providers. From birth, caregivers form close and collaborative relationships with their PCPs. These relationships and the frequency of visits with the PCP allow for close and consistent monitoring of developmental progress ranging from routine hearing and vision checks to developmental milestones (reflexes, social skills, motoric development). At approximately 9 to 18 months of age, and even earlier in some cases (e.g., present risk factors including children with history of preterm birth, low birth weight, or sibling with autism spectrum disorder), PCPs incorporate screening practices to identify developmental delays. The American Academy of Pediatrics supports universal autism screening, recommending screening at children’s 18- and 24-month well visits, as research has shown that screening leads to earlier referral and diagnosis. Children who are screened earlier, receive services earlier, which leads to better outcomes (McPheeters et al., 2016). Thus, PCPs offer a unique relationship in which on-
SCREENING AND ASSESSMENT

going assessment and monitoring of challenging behavior can occur, a hallmark of an MTSS framework.

McGuire and colleagues (2016) first recommended to screen for the presence of challenging behavior. Next, if challenging behavior is found to occur, safety is assessed. That is, it should be determined if the individual is at imminent risk of injuring themselves or others. At this juncture, if challenging behavior that presents a safety risk is occurring, and the resources are available, the PCP should refer the individual to a more specialized provider with experience in assessment and treatment of challenging behavior. If a safety risk is not present, or more specialized resources are not available, then the PCP can review the patient’s psychosocial history and level of functioning before and after the onset of challenging behavior. This would include assessment of the individual’s medical, developmental, communicative, and psychiatric history, and may require referrals to specialists with expertise in these areas. Further, assessment of the individual’s current environment, those providing care for the individual, and the individual’s adaptive functioning are recommended. Fourth, challenging behavior should be prioritized based on safety, severity, and the impact that it has on the individual’s daily life. From there, all potential contributors to challenging behavior should be considered, such as: underlying medical problems (e.g., pain, seizures, gastrointestinal issues), functional communication difficulties, psychosocial stressors, maladaptive reinforcement patterns, and co-occurring psychiatric disorders. Then, based on these outcomes, referrals should be made to specialized providers to coordinate an individualized treatment plan. Finally, the treatment plan should be implemented, by the appropriate providers, and monitored for effectiveness at 3-month follow-up and every 3 months thereafter. If challenging behavior improves, less intense
SCREENING AND ASSESSMENT
intervention may be considered; if challenging behavior worsens, a higher level of care may be
needed. Below, a schematic of this pathway is provided in Figure 5.
Figure 5

Primary-Care pathway

Assess for challenging behavior

Unsafe

Assess for safety

Safe

Consider transfer to higher-level care

Review the patient’s history and level of functioning

Prioritize and qualify challenging behaviors for treatment

Consider potential contributors to challenging behavior

Current medical problems
Functional communication difficulties
Psychosocial stressors
Maladaptive reinforcement patterns
Co-occurring psychiatric disorders

Consider psychopharmological interventions for challenging behavior

Coordinate individualized treatment and safety plans

Monitor individualized treatment and safety plans

Yes

At 3 months do symptoms persist?

No

Re-evaluate every 3 months thereafter

SCREENING AND ASSESSMENT

*Progress Monitoring*

After challenging behavior has been identified, and appropriate services have been accessed, the effects of treatment need to be monitored such that if approaches are showing to be ineffective, alternative treatment modalities can be explored before challenging behavior worsens. Direct observation of challenging behavior is a highly sensitive measurement strategy that can detect subtle changes in challenging behavior over time, making it ideal for progress monitoring. Strategies for directly measuring challenging behavior are discussed at length earlier in this document, and although these tactics are good for monitoring progress, they are extremely resource intensive. It is not feasible, for example, to directly observe an individual’s daily behavior over months and years. Therefore, alternative measurement strategies that are sensitive to change, but less resource intensive are needed.

The electronic medical records (EMR) system is also an important source to capture and monitor progress. The EMR allows for reviewing and visualizing changes in the child’s developmental profile. It also promotes consistent measurement, as providers will be more likely to utilize the same measure when it has been used previously and its contents are built into the EMR. Without consistent utilization of the same measure, it is hard to understand changes over time. Use of the EMR allows for pre-programmed alerts, which remind the provider to screen on a routine basis.

There are, however, times when a child’s behaviors change rapidly or stakeholders are unable to catch the red flags prior to the behaviors becoming particularly impactful or harmful (i.e., no history of less intense responses occur prior to behavioral escalation). In these cases, the frequency and intensity of the behavioral incident may warrant emergent (e.g., emergency room visit, crisis center) versus planned or programmed action (e.g., behavioral surveillance, routine
SCREENING AND ASSESSMENT
check-ups, formative/summative academic meeting) from caregivers and practitioners.

Unfortunately, even isolated instances of severe challenging behavior can have deep and lasting impact for the youth, family, and community warranting immediate intervention. Practitioners may employ additional interviews, checklists, questionnaires, and rating scales that are more specifically targeted at better understanding the level, frequency, and intensity of the presenting concerns in relation to same-aged peers. In addition, practitioners may recommend more intensive progress monitoring on a denser schedule. Potential adverse events associated with the assessment and treatment of challenging behavior should also be measured to ensure the individual is receiving appropriate care. Future research is needed to develop sensitive and standardized measures that could specifically be used for monitoring progress over time and further differentiating acute versus chronic episodes of challenging behavior.

Summary

Overall, an organized network for screening and surveilling youth with ASD for development of challenging behavior does not exist and the current procedures are inadequate. It is critical to prevent severe challenging behavior from emerging that routine well-child checkups (early childhood) and annual well-checks (later childhood adolescents, adulthood) occur and include screening and follow-up for challenging behaviors. During these visits, challenging behaviors need to be discussed and brought to the attention of the child’s PCP. Similarly, assessment for development of socially appropriate behaviors should also be cataloged such that caregivers and practitioners actively work towards skill development when gaps or delays are identified (i.e., unlearning the “wait it out” model). Ideally, screeners should then be employed and used to inform the referral to appropriate service providers (see Primary-Care Pathway). Indirect measures (e.g., structured and unstructured interviews, rating scales, questionnaires;
SCREENING AND ASSESSMENT
specific tools are detailed below) are often a first choice given ease of implementation and lower
cost, relative to direct and experimental methods of assessment. That being said, the benefit of
low-resource intensiveness and accessibility also comes with a drawback in the areas of accuracy
and reliability (Iwata et al., 2000). Unlike direct assessment, indirect assessments do not require
direct observation of the patient. Thus, when considering the use of indirect assessments,
practitioners should consider employing these measures in combination with other direct or
structured observation methods to optimize reliability and accuracy. Since many practitioners
may only observe patients in the context of routine well-child check-ups, we strongly
recommend that practitioners ask caregivers to support indirect measures with samples of the
child’s behavior by providing video samples or pictures whenever possible (Iwata & DeLeon,
1996). On-going assessment should then be employed to monitor progress and direct future
referrals if needed.

Conclusion

Caregivers, pediatricians, community mental-health providers, and school personnel
should all engage in a collaborative effort to identify deviations from developmentally normative
and clinically significant levels of challenging behavior in children with ASD. Initially,
stakeholders in an individual’s care (e.g., PCP, school psychologist) screen for challenging
behavior by employing a combination of indirect and direct assessment strategies to gather
information about a child’s developmental presentation and presence of challenging behavior in
relation to same-aged peers. These assessment strategies identify concerns related to the form,
frequency, duration, and/or intensity of challenging behaviors, as well as consideration of the
individual’s strengths and preferences. The information gathered from these assessments
illuminate next steps for clinical decision making. Then, FBA may be needed to directly examine
SCREENING AND ASSESSMENT
the environmental influences on challenging behavior, which in turn is prescriptive for treatment (see Treatment section).

By continuing to monitor challenging behavior over time, perhaps during routine PCP visits, individuals can be linked to the assessment and treatment approaches that meet their, and their families, unique and ever-changing needs. For example, for more severe and frequent forms of challenging behavior, specialized settings that provide psychiatric assessment and functional analysis methodologies to inform treatment may be recommended (e.g., inpatient care). Over time, however, adaptive functioning may improve, and outpatient or community-based care may be better suited and preferred. Repeated measurement of challenging behavior over time and across services is needed to match treatment to each individual’s needs and strengths. In closing, challenging behavior is prevalent in the ASD population and associated with a host of negative health outcomes. Early identification and assessment are needed to provide earlier intervention and improve quality of life.
SCREENING AND ASSESSMENT

References


SCREENING AND ASSESSMENT


SCREENING AND ASSESSMENT


SCREENING AND ASSESSMENT


SCREENING AND ASSESSMENT


SCREENING AND ASSESSMENT
Autism and Developmental Disorders, 48(11), 3678–3688. APA PsycInfo.
https://doi.org/10.1007/s10803-017-3460-2


SCREENING AND ASSESSMENT


SCREENING AND ASSESSMENT


SCREENING AND ASSESSMENT


SCREENING AND ASSESSMENT


SCREENING AND ASSESSMENT


SCREENING AND ASSESSMENT


SCREENING AND ASSESSMENT


SCREENING AND ASSESSMENT
https://doi.org/10.1901/jaba.1995.28-91


https://doi.org/10.1002/aur.2224


In J. L. Matson (Ed), Functional assessment for challenging behaviors Springer, New York, NY.


SCREENING AND ASSESSMENT


