behavioral sciences

MDPI

Review

A Critical Appraisal of the Measurement of Adaptive Social Communication Behaviors in the Behavioral Intervention Context

Thomas W. Frazier [1,2,3,*], Eric A. Youngstrom [4], Allison R. Frazier [5] and Mirko Uljarevic [6]

[1] Department of Psychology, John Carroll University, 1 John Carroll Blvd, University Heights, OH 44118, USA
[2] Departments of Pediatrics and Psychiatry, SUNY Upstate Medical University, Syracuse, NY 13210, USA
[3] Autism Speaks, Princeton, NJ 08540, USA
[4] Institute for Mental and Behavioral Health Research and the Department of Psychiatry, Nationwide Children's Hospital and The Ohio State University, Columbus, OH 43215, USA; eric.youngstrom@nationwidechildrens.org
[5] Elevate Learning, LLC, Pepper Pike, OH 44124, USA; allison@ohelevate.com
[6] Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford, CA 94305, USA; mirkoulj@stanford.edu
* Correspondence: tfrazier@jcu.edu

Abstract: Despite encouraging evidence for the efficacy of comprehensive and intensive behavioral intervention (CIBI) programs, the majority of studies have focused on relatively narrow, deficit-focused outcomes. More specifically, although adaptive social communication and interaction (SCI) are essential for facilitative functioning, the majority of studies have utilized instruments that capture only the severity of SCI symptoms. Thus, given the importance of the comprehensive and appropriate characterization of distinct SCI adaptive skills in CIBI, in this review, based on PubMed search strategies to identify relevant published articles, we provide a critical appraisal of two of the most commonly used adaptive functioning measures—the Vineland Adaptive Behavior Scales-Third Edition (Vineland-3) and the Adaptive Behavior Assessment System-Third Edition (ABAS-3), for characterizing SCI in the behavioral intervention context. The review focused on periodic outcome and treatment planning assessment in people with autism spectrum disorder receiving CIBI programs. Instrument technical manuals were reviewed and a PubMed search was used to identify published manuscripts, with relevance to Vineland-3 and ABAS-3 development, psychometric properties, or measure interpretation. Instrument analysis begins by introducing the roles of periodic outcome assessment for CIBI programs. Next, the Vineland-3 and ABAS-3 are evaluated in terms of their development processes, psychometric characteristics, and the practical aspects of their implementation. Examination of psychometric evidence for each measure demonstrated that the evidence for several key psychometric characteristics is either unavailable or suggests less-than-desirable properties. Evaluation of practical considerations for implementation revealed weaknesses in ongoing intervention monitoring and clinical decision support. The Vineland-3 and ABAS-3 have significant strengths for cross-sectional outpatient mental health assessment, particularly as related to the identification of intellectual disability, but also substantial weaknesses relevant to their application in CIBI outcome assessment. Alternative approaches are offered, including adopting measures specifically developed for the CIBI context.

Keywords: autism spectrum disorder; behavioral intervention; Vineland-3; ABAS-3; adaptive behavior

check for updates

Academic Editor: Dario Siniscalco

Received: 10 March 2025
Revised: 15 May 2025
Accepted: 20 May 2025
Published: 23 May 2025

Citation: Frazier, T. W., Youngstrom, E. A., Frazier, A. R., & Uljarevic, M. (2025). A Critical Appraisal of the Measurement of Adaptive Social Communication Behaviors in the Behavioral Intervention Context. *Behavioral Sciences*, *15*(6), 722. https://doi.org/10.3390/bs15060722

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Over the last several decades, a range of targeted practices that use the principles and practices of behaviorism and more comprehensive intervention packages that typically integrate multiple specific practices with an over-arching discipline or philosophy (e.g., applied behavior analysis or developmental psychology) have been developed and evaluated in youth with autism spectrum disorder (ASD) (Hume et al., 2021; Trembath et al., 2023). Despite encouraging evidence for the efficacy of the noted approaches, in particular, with regard to improvements in IQ and language abilities, the majority of treatment trials have focused on a relatively narrow range of outcomes, emphasizing the need to develop assessments capturing behaviors and skills that are relevant to a person's ability to function across different daily contexts. Further, over the last decade, both theoretical frameworks and dimensional initiatives, such as the National Institute of Mental Health's Research Criteria, and latent variable modeling studies, have emphasized the multi-dimensional nature of specific aspects of adaptive functioning, in particular, the social communication and interaction domain, demonstrating that different subdomains have distinct trajectories and underpinning mechanisms and might thus respond differently to specific treatments. Therefore, by providing only overly broad scores that conflate distinct subdomains, outcome measures can obscure the potential positive effects of specific treatments. In addition to the importance of domain coverage and representation, modern psychometric approaches have made significant strides towards regression-based norming and other approaches that can be utilized to optimize instruments to inform treatment selection as well as to quantify treatment-related change at the individual patient level and thus significantly improve the science and practice of behavioral interventions. Crucially, with the recognition of the importance of neurodiversity-affirming practices, it has become increasingly clear that it is essential to go beyond simply focusing on symptom-/deficit-based outcomes and approaches (Schuck et al., 2022; Sulek et al., 2024; Waddington et al., 2024). Given these recent developments, it is essential to ensure that outcome measures are psychometrically robust and provide a comprehensive capture of all the relevant adaptive subdomains without conflating distinct constructs. Further, it is important to establish whether currently used instruments have utilized state-of-the-art psychometrics to provide modern norming, the ability to track change, and facilitate practice through automated online administration, scoring, and interpretation. Thus, in this review, we focused on appraising whether two of the most widely used adaptive functioning instruments—the Vineland Adaptive Behavior Scales-Third Edition (Vineland-3) and the Adaptive Behavior Assessment System-Third Edition (ABAS-3), provide construct coverage in line with the current empirical evidence and incorporate psychometric advances to enable their valid use as outcome measures for behavioral treatments.

1.1. Types of CIBI Packages

The most common comprehensive and intensive behavioral intervention (CIBI) packages for ASD can be grouped into a more structured or traditional application of applied behavior analysis (ABA) that includes early intensive behavioral intervention approaches that utilize a more directive and structured approach to intervention (Cooper et al., 2007) and naturalistic developmental behavioral interventions (NDBIs) that emphasize following the child's lead, using naturalistic reinforcement in everyday settings to promote developmentally appropriate skills (Schreibman et al., 2015). Importantly, even in research applications of ABA and NDBI, clinicians often modify the application of the intervention package to the characteristics of the child, potentially even borrowing from features of the other package in implementation (Rogers et al., 2021). For instance, in situations where the need for stimulus and environmental control is high, naturalistic delivery is less crucial or

relevant, and/or child-led intervention is less likely to be effective, many ABA practitioners, including Board Certified Behavior Analysts, need to adapt their intervention delivery methods to the specific profile of individual's characteristics, using aspects of ABA with strong developmental consideration. In addition, it is essential for practitioners to consider how meaningful and valid specific treatment goals are to the given family and individual (Sulek et al., 2024; Waddington et al., 2024). Regardless of individual perspectives on the ABA and NDBI approaches, a growing number of individuals are accessing CIBI packages, and, as a result, there is a need for useful and appropriate outcome measures to inform the process and evaluate the value of these treatments and to consider a reduction in specific symptoms and, equally importantly, the development of specific aspects of adaptive functioning and quality of life (Waddington et al., 2024).

1.2. Assessment of Social Communication/Interaction Skills in CIBI

Although a major focus of CIBI has, historically, been on reducing the severity of social communication and interaction (SCI) symptoms, it has also been suggested that it is crucial for treatments to focus on the development and improvements in the specific SCI adaptive skills that underpin one's ability to navigate the complexity of the social world, and, if impaired, result in the symptoms that characterize ASD. However, one of the major issues in terms of evaluating the effects of the specific treatments has been the fact that the most widely used instruments often conflate skills and symptoms, and adaptive SCI behaviors are often the other side of the autism symptom coin, with SCI skills being replacement behaviors for core autism symptoms. The highly similar nature of SCI measurement, when conceptualized as an adaptive skill or as a symptom, is apparent when comparing social communication and interaction items across adaptive functioning and autism symptom measures. For instance, the following item stem used to measure the socialization domain on the Vineland-3 "Realizes when others are happy, sad, . . ." is highly similar to the following item measuring perspective taking on the Autism Symptom Dimensions Questionnaire "Seem to understand what others are thinking or feeling". Strong item similarity is not an exception, as many adaptive behavior items assessing SCI behaviors overlap significantly with social communication content on autism symptom measures (see Table S1).

Given the prominence of SCI behaviors as intervention targets for individuals with ASD, including their important contribution to more positive long-term outcomes (Forbes et al., 2023), the present paper focuses on the potential of adaptive function measures as measures of SCI within CIBI outcome assessment instruments. Specifically, we first briefly review historical and research approaches to evaluating the benefits and outcomes of CIBI. Then, the potential roles of and needs for outcome assessment in CIBI are covered. Next, modern conceptualizations of SCI behavior are reviewed to set the stage for evaluation of two of the most commonly administered adaptive functioning measures, the Vineland Adaptive Behavior Scales-Third Edition [Vineland-3] (Sparrow et al., 2016) and Adaptive Behavioral Assessment System [ABAS-3] (Harrison & Oakland, 2015), as measures of SCI behavior. Evaluation of these instruments as measures of SCI for CIBI includes a review of how the measures were developed, current evidence supporting their development, psychometric characteristics (including factor structure, measurement invariance, reliability, and validity), and practical aspects of their implementation to inform CIBI practice. Finally, this evaluative review ends by discussing how the field can shift toward more optimized approaches to adaptive behavior outcome assessment as part of a larger strategy for assessing CIBI outcomes and informing care for people with ASD.

1.3. Historical Approaches to Outcome Assessment for CIBI

As different treatment models, including the NDBI and ABA models, were adopted by specialty care centers and payor coverage became more widespread, the need for intermittent assessment of cognition and behavior as a means of evaluating and demonstrating the value of the everyday community implementation of CIBI increased (Smith et al., 2015). The batteries of measures implemented often depended on the specific treatment model or philosophy, including a mixture of standardized and unstandardized and proximal (e.g., measurement of behaviors as they occur during treatment) versus distal (e.g., assessment of cognitive or behavioral constructs at given timepoints during the intervention course) measurement strategies. In general, periodic outcome assessments frequently included instruments assessing general cognitive ability, language, speech/communication, adaptive function, and autism symptoms (Ridout & Eldevik, 2023). A standardized, robust, well-validated, comprehensive set of assessments capturing all aspects of the clinical phenotype as well as other key cognitive and adaptive skills is essential to ensure that every patient is assigned to appropriate treatments and that their responses to interventions are sensitively monitored to enable changes in clinical management. However, substantial variability has been present within and across clinical and research contexts. This variability has even extended across payor policies for behavioral intervention delivery models. More specifically, while some payors are adopting specific measure sets that are required to be collected at regular intervals (often every 6–12 months), such as TRICARE (https://www.tricare.mil/autism (accessed on 1 August 2024)), other payors and trade organizations are providing general guidance for measure selection or leaning on existing best practices and recommendations (https://www.casproviders.org/standards-and-guidelines (accessed on 1 August 2024)). Within the last 18 months, organizations have even begun to recommend autism-specific outcome batteries (Joseph et al., 2024).

1.4. Current Challenges in CIBI Outcome Assessment

Given the current state of the assessment field and the sometimes poorly specified or capricious nature of requirements or recommendations within and across funder policies, several initiatives have attempted to address these issues, with significant impact on both clinical practice and policy. For example, the National Academies of Science Engineering and Medicine (NASEM) was asked by Congress to evaluate the Autism Care Demonstration (ACD) project within the TRICARE military health benefit. Part of the ongoing analysis being performed by NASEM is expected to include making recommendations regarding future outcome assessment strategies (https://www.nationalacademies.org/our-work/independent-analysis-of-department-of-defenses-comprehensive-autism-care-demonstration-program (accessed on 1 August 2024)). Through the process adopted by NASEM, specific concerns have been raised about whether TRICARE ACD legacy assessment approaches are appropriate for outcome assessment of CIBI, whether measures are covering meaningful domains of functioning relevant to patients and families, and practical concerns regarding the feasibility of measure completion and whether measures are being appropriately scored and utilized to inform intervention.

Thus, there is a crucial need as well as an opportunity to develop and validate more rational, evidence-based, or evidence-informed outcome assessment practices for CIBI in ASD. This need is occurring against a backdrop of significant advances in psychological assessment, including the increasing adoption of recommendations for measure development practices (Boateng et al., 2018; FDA, 2009) as well as technical advances in how measures are psychometrically evaluated and practically delivered. More specifically, over the last decade, there have been substantial improvements in the evaluation of measure structure (Marsh et al., 2014), norming (Sorensen et al., 2023), scoring for longitu-

dinal monitoring (C. A. Farmer et al., 2020), assessment of reliable and meaningful change (Chatham et al., 2018), and practical improvements in instrument deployment such as automated online administration, scoring, reporting, monitoring, and the provision of decision support (Clausen et al., 2021). Unfortunately, the process of the development, validation, and deployment of the most widely used outcome assessment of CIBI has not incorporated noted psychometric advances that are essential for delivering instruments fully optimized for comprehensive characterization, score interpretation, and treatment monitoring.

In light of the current circumstances, and given the nascent nature and growing availability of CIBI for ASD and the fact that a substantial proportion of all BCBAs are less than 5 years from certification (https://www.bacb.com/bacb-certificant-data/ (accessed on 1 August 2024)), there is opportunity and motivation to re-think CIBI outcome assessment at the instrument and process levels. This is particularly crucial in light of recent discussions in the field emphasizing the need to incorporate perspectives from the community and focus on outcomes identified by autistic individuals and their families as being important, meaningful, and socially valid (Baiden et al., 2024; Schuck et al., 2024a; Schuck et al., 2024b). Thus, this review focuses on understanding the characteristics and potential utility of existing adaptive function measures within CIBI outcome assessment, with an eye toward what improvements might be possible with additional advances in measure development, evaluation, and implementation.

1.5. Multiple Roles for Assessment Processes Within CIBI

In 2024, the Council of Autism Service Providers published The Applied Behavior Analysis Practice Guidelines for the Treatment of Autism Spectrum Disorder Guidance for Healthcare Funders, Regulatory Bodies, Service Providers, and Consumers (Council of Autism Service Providers, 2024), which includes a description of initial and periodic assessment activities as well as considerations for monitoring and adjusting treatment in order to achieve optimal outcomes. After a diagnosis of ASD, and a referral for CIBI services, generally accepted standards of care include using a multi-modal assessment approach to inform the CIBI treatment plan and ongoing clinical management (Figure S1). CIBI assessment methods can be broadly categorized into three types: (1) session-based data collection on specific intervention targets relevant to pre-specified objectives, (2) the collection of intervention quality and fidelity data, and (3) periodic outcome and treatment planning assessment. All three assessment types inform the provision of services, including addressing below-expectation skills and reducing or replacing maladaptive behaviors. Each assessment type adds distinct value to CIBI, but the present review focuses on periodic CIBI outcome and treatment planning assessments. For additional information on session-based and fidelity/quality assessment see Supplemental Background.

A rigorous, periodic (often every 6 months) CIBI outcome and treatment planning assessment process (Figure S2) is often narrowly conceptualized as determining the strength of the individual's intervention response at a given point in time, enabling stakeholders (including payors or health plans) to identify patient progress and inform future clinical management, including authorizations of services for additional treatment periods. While this is an important aspect of CIBI outcome assessment, it is not the only value that the application of outcome measures can provide. Instead, and consistent with existing guidelines (Council of Autism Service Providers, 2024), this review posits at least five major roles that initial and ongoing CIBI outcome assessment can fulfill: (1) developing the initial CIBI strategy, (2) identifying cognitive and behavioral domains and sub-domains, (3) recommending specific behaviors to address within treatment protocols, (4) monitoring intervention progress and determining the strength of intervention response, and (5) highlighting possible intervention modifications and changes in clinical management. It is

crucial to consider these roles when evaluating the Vineland-3 and ABAS-3 as measures of SCI adaptive functioning, and to understand how future measurement instruments, including revisions of these instruments, could provide additional value. In addition, it is important to consider the recent criticisms of the exclusive deficit-focused orientation of the more traditional CIBI approaches and the lack of alignment between traditionally defined intervention outcomes and the outcomes that were identified as most meaningful by autistic individuals and their families, including outcomes that are related to improvements in different aspects of quality of life (Baiden et al., 2024; Sulek et al., 2024). Supplemental Text S1 provides detailed information about specific roles for periodic CIBI outcome and treatment planning assessments.

1.6. Modern Understanding of SCI Structure and Relevant Assessment Considerations

There have been substantial developments in the understanding of the structure and taxonomy of specific facets of adaptive social communication and interaction behavior. For instance, both theoretical frameworks (Happe et al., 2017; Huber et al., 2019) and single- and multi-instrument factorial work (Chetcuti et al., 2024; Frazier & Hardan, 2017; Phillips et al., 2019; Simmons et al., 2024; Uljarevic et al., 2020a) have highlighted the multifaceted nature of social and communication functioning, identifying a specific set of processes underpinning one's ability to adapt to the complexities of daily life that distinct and often non-linear development progression. Specifically, detailed, large-sample factor-analytic work using neurotypical, ASD, and other neurodevelopmental disability samples has identified an overarching social communication and interaction behavior domain (in contrast to separate domains for social interaction and communication) and at least four subdomains of social communication and interaction behavior (see Table S2). The identified sub-domains include social motivation or affiliation-related behaviors, basic social communication behaviors (including verbal and non-verbal aspects), perspective-taking or theory of mind-related behaviors (comprising both cognitive and emotional aspects), and the frequency and quality of relationship-focused and reciprocal interaction behaviors. Importantly, the noted domains and sub-domains emerged across the factor analyses of both instruments that capture SCI abilities/functioning (e.g., Stanford Social Dimensions Scale (Phillips et al., 2019) and SCI-related symptoms (e.g., Social Responsiveness Scale, Social Communication Questionnaire, and Autism Diagnostic Interview-Revised) (Frazier et al., 2008, 2012, 2014, 2023a, 2023b; Frazier & Hardan, 2017; Georgiades et al., 2012; Mandy et al., 2012; Shuster et al., 2014; Snow et al., 2009; Uljarevic et al., 2020a, 2021). These domains are also aligned with current transdiagnostic frameworks, including the National Institute of Mental Health's Research Domain Criteria (Insel et al., 2010).

Given the complex and multifaceted nature of the SCI domain, it is crucial that CIBI outcome assessment batteries include SCI domain coverage and representation that is aligned with noted conceptual and empirical advancements to ensure the adequate measurement of the full breadth of SCI behaviors. If specific SCI dimensions are not explicitly assessed or are only partially measured, it will be difficult to build and monitor an appropriate intervention plan. For example, in a child with lower social motivation and weaker basic social communication skills but with a recently developed friend, clinicians may be tempted to focus heavily on perspective-taking and relationship skills but would be better served by addressing the more basic social motivation and basic social communication skills before moving toward more complex and higher-level social cognitive skills. The principle of strong sub-domain coverage has particular relevance for adaptive function measures, which were not developed with core ASD symptom dimensions or a modern understanding of SCI behavior and, therefore, tend to focus on the very broad measurement of the SCI construct. Reinforcing the importance of capturing distinct aspects of SCI are recent studies

that have demonstrated the utility of fine-grained SCI subdomains for characterizing ASD heterogeneity and identifying informative subgroups. For instance, a relatively recent study utilizing a comprehensive SCI measure has identified five subgroups of youth with ASD that showed distinct patterns of strengths and weaknesses across different aspects of the social motivation, basic social communication skill, and theory of mind SCI subdomains rather than simply reflecting a severity gradient and, importantly, the identified subgroups were further differentiated in terms of cognitive ability, the severity of ASD symptoms, and the severity of co-occurring internalizing symptoms (Uljarevic et al., 2020b).

Evidence-based assessment frameworks point out that the clinical goals of diagnosis and treatment planning versus progress, process, and outcome evaluation not only occur at different phases in the arc of a course of treatment, but they also often have distinct features that can sometimes be in tension (Youngstrom & Prinstein, 2020). Given that adaptive function instruments were not specifically developed with CIBI in mind, but are nevertheless widely deployed in the CIBI outcome assessment process, it is essential to examine how these measures might inform specific clinical activities in the context of the necessary roles fulfilled by the periodic CIBI outcome assessment process (Supplemental Text S1). Additionally, as part of evaluating the Vineland-3 and ABAS-3 in terms of their ability to measure SCI behavior and fulfill the above-described roles within periodic CIBI outcome assessment, it is useful to consider their development processes and, in particular, the lack of alignment with the modern conceptual and empirical understanding of the SCI domain, psychometric characteristics, and practical considerations in their deployment. Thus, before providing a detailed evaluation of these instruments, we first outline the measure development process and the psychometric and practical features needed for any instrument contributing to effective CIBI outcome measurement. Then, within each section, we focus on evaluating these considerations for the measurement of SCI by the Vineland-3 and ABAS-3.

2. Review Methods and Results

For the present critical review, two literature searches were conducted using PubMed. For the Vineland-3, search terms included Vineland AND adaptive AND (psychometric OR reliability OR validity OR development). For the ABAS-3, search terms included ABAS OR adaptive behavior assessment system AND adaptive AND (psychometric OR reliability OR validity OR development). The PubMed searches generated 749 results for the Vineland and 1822 results for the ABAS. The titles and abstracts were reviewed for relevance, and the reference lists were reviewed for articles with strong relevance to identify any missed publications. The Vineland-3 and ABAS-3 technical manuals were also reviewed for relevant information.

Over the last 15 years, several efforts have been made to provide guidance to measure developers regarding best practices in instrument development. In particular, the FDA has provided substantial guidance regarding patient-reported or observer-reported outcome measure development (FDA, 2009). Similarly, measure development experts have recommended sets of procedures and processes that can be used to inform measure development and subsequent psychometric evaluation and validation (Boateng et al., 2018). Table 1 presents the measure development guidelines, based on expert recommendations, that are most relevant to application in CIBI and their relationship to Vineland-3 and ABAS-3 development. These guidelines were used to evaluate the Vineland-3 and ABAS-3 development and validation processes. Supplemental Methods provide additional information on the literature search.

Table 1. Measure development guidelines most relevant to the CIBI outcome assessment process and the evaluation of the Vineland-3 and ABAS-3 across these guidelines.

		Vineland-3	ABAS-3
1	Stakeholder involvement from conceptualization to validation	Partial—prior administrators, clinicians, and clinical researchers	Partial—experts, literature review with item refinement including stakeholders
2	Use of qualitative research processes, including concept elicitation interviews, in determining domains/sub-domains guiding measurement coverage	Not present	Not Present
3	Creation of a domain map or disease model to guide measure creation	Not present	Present for IDD, not specific to ABA
4	Consideration of CIBI population-specific issues	Partial—IDD population	Partial—IDD population
5	Attention to the nature of anticipated measure utilization (e.g., diagnostic or outcome measurement)	Not present	Not present
6	Use of cognitive interviewing to evaluate item appropriateness	Partial—for teacher-reported versions	Unknown
7	Assessment of item relevance, clarity/readability, and potential bias	Present—review for bias and relevance/importance	Present—review for clarity/relevance

2.1. Measure Development Processes

Both the Vineland-3 (Sparrow et al., 2016) and ABAS-3 (Harrison & Oakland, 2015) include stakeholders with developmental disabilities in the measure development process and report in their manuals methods which appear to have involved evaluation of item characteristics, including relevance, clarity/readability, and potential biased language. However, neither instrument's manual description includes mention of parent/caregiver or patient/client stakeholder inclusion at the outset of conceptualization, using concept elicitation or cognitive interviews, or creating a domain map with stakeholders. From the manuals, it does not appear that much of the development work occurred within a behavioral intervention context, as the primary initial development rationale emphasized the measurement of adaptive behavior as part of the identification of intellectual and developmental disability (IDD). Furthermore, it is apparent that the full developmental disability population was not considered, as, for example, the Vineland-3 sub-domains and items (e.g., Communication sub-domain: Reading and Writing) have very limited relevance to those with significant cognitive challenges or profound autism (Lord et al., 2021). As a result, neither the Vineland-3 nor ABAS-3 were purpose-built for the CIBI outcome assessment context.

2.2. Psychometric Considerations

The psychometric properties most important to CIBI outcome assessment are listed in Table 2, including an overview of existing evidence for the Vineland-3 and ABAS-3 relevant to each characteristic. While a comprehensive psychometric evaluation of any psychological assessment instrument can involve a wide range of methods and characteristics, the present review focuses on those most crucial for the assessment of social communication and interaction behavior within a CIBI context.

Table 2. Psychometric criteria for evaluation of Vineland-3 and ABAS-3 use in CIBI outcome assessment.

	Criterion	Vineland-3	ABAS-3
		Structure	
1	Factor structure	Inconsistent findings for subdomains across publications, no item-level analysis	Inconsistent findings, no item-level analysis
2	Measurement invariance/differential item or scale functioning	Inconsistent evidence for clinical IDD groups but good invariance evidence for sex, race/ethnicity, and SES	Invariance evidence for age and sex/gender
3	Measurement model guides scoring	Scoring based on theoretical, not empirical considerations	Scoring based on theoretical, not empirical considerations
		Reliability	
4	Scale reliability	Very strong for scored scales	Very strong for scored scales
5	Conditional reliability	Not evaluated	Not evaluated
6	Test–retest reproducibility	Very good	Very good
7	Test–retest stability	Not evaluated	Not evaluated
8	Inter-rater reliability	Good to excellent for raters within forms, across interviewers	Good to excellent for raters across forms with overlapping ages, good across adult self and adult informant reports
		Validity	
9	Content coverage	Good to excellent for broad SCI domain, weaker for specific SCI content areas	Good to excellent for broad SCI domain, weaker for specific SCI content areas
10	Construct coverage	Limited evidence to support domain sub-dimensions, no coverage of key SCI subdomains	Limited evidence to support domain sub-dimensions, no coverage of key SCI subdomains
11	Convergent validity	Strong evidence for broad domain scores across numerous studies, including criterion-related validity in ASD samples; small to medium correlations between Vineland domains/sub-domains and ASD instrument scores; correlations with ASD symptom dimensions are similar across communication, daily living, and social domains	Good evidence for broad domain scores across several studies, including criterion-related validity in ASD samples, correlations with ASD symptom measures are moderate and comparable in size across the three SCQ domains for SCQ; some evidence of convergent validity with ADOS-2
12	Discriminant validity	Excellent	Excellent
13	Sensitivity to change	Evidence for significant and clinically meaningful gains across observational and randomized studies for socialization domain score	Not evaluated
		Norming	
14	Sample representativeness	Strong representativeness	Strong representativeness
15	Appropriate demographic adjustment	Yes, age adjustment for standard scores	Yes, age adjustment for standard scores
16	Traditional vs. continuous norming	Mixed norming, with accounting for distribution skewness and non-linear age trends in v-scale scoring	Traditional norming, but with a very large sample ~4000 and matching to US census

Table 2. *Cont.*

	Criterion	Vineland-3	ABAS-3
		Scoring/Reporting	
17	Validity indicators	Not present	Not present
18	Raw scores	Present by domain and subscale	Present by domain and skill area
19	Norm-referenced scores	Present by domain and subscale	Present by domain and skill area
		Change Measurement	
20	Growth scores	Present by domain	Not present
21	Reliable change scores	Not present	Not present
22	Clinically meaningful change	Published levels available	Not present

Note. SCQ = Social Communication Questionnaire. ADOS-2 = Autism Diagnostic Observation Schedule-Second Edition.

2.2.1. Overview of Instrument Structure and Scoring

The Vineland-3 (Sparrow et al., 2016) is scored with up to 4 domain scores, including separate socialization and communication domains. The socialization and communication domains include 6 subscale scores (Socialization: Interpersonal Relationships, Play and Leisure, Coping Skills; Communication: Receptive, Expressive, Written). A 3-point Likert frequency scale is used for each item that corresponds to "Never", "Sometimes", and "Usually/Often". The ABAS-3 (Harrison & Oakland, 2015) has two broad domains and three skill areas relevant to SCI behavior (Conceptual: Communication and Social: Social and Leisure). The items use a 4-point Likert frequency scale—"Never", "Sometimes", "Often", and "Very Often". In addition to the parent/caregiver report versions, the Vineland-3 offers an interview form and both the Vineland-3 and ABAS-3 offer teacher report versions. The teacher report versions are less relevant to a CIBI context where additional raters are often not available or obtaining reliable information from them is not feasible. The interview version is recommended in the Vineland-3 manual (Sparrow et al., 2016), with the implication that introducing clinical judgment yields more accurate or valid measurement. However, this review did not identify any peer-reviewed publications that specifically demonstrate greater validity for the interview version of the Vineland-3 (see Supplemental Text S2 for additional information). While clinical judgment has a longstanding and well-worn history of being used to integrate information and clarify diagnostic and phenotypic presentations, there are numerous examples where clinical judgment actually introduces substantial variability across interviewers or sites (Rettew et al., 2009), a significant problem for real-world CIBI programs where staff turnover can be frequent. While it is certainly possible that clinical judgment does enhance accuracy or validity, the absence of evidence, particularly with regard to the specific contexts in which clinical judgment might provide additional insights, raises significant concerns as to whether the significant cost, clinician time, and burden for families related to the use of the interview form of the Vineland-3 are justified. Further, this represents a major challenge for some clinics where providers may not have sufficient training in clinical interviewing.

2.2.2. Factor Structure and Measurement Invariance

Although often overlooked, before considering measure scoring, it is critical that the structure of the instrument be understood. Recent developments in latent structural analysis permit the detailed comparison of competing models, with best-fitting models being used to inform instrument scoring. Unfortunately, the factor-analytic literature of the Vineland-3 and ABAS-3 is very limited, both in terms of the number of structural evaluations but also regarding the use of a limited range of methods for identifying optimal struc-

ture. Of the published factor-analytic evaluations of the Vineland-3 (de Bildt et al., 2005a; R. L. Farmer et al., 2021; Pandolfi & Magyar, 2021; Wilkinson et al., 2023), the analyses used only subscales as indicators (not items) and the results were inconsistent, suggesting from one to three factors but often with poor or adequate but not excellent model fit. For example, two of the identified Vineland-3 confirmatory factor-analytic studies reported poor fit (Pandolfi & Magyar, 2021; Wilkinson et al., 2023), below the recommended levels for adequate or excellent comparative fit indices (Marsh et al., 2004). Furthermore, measurement invariance and differential item functioning analyses suggested potential measurement inconsistencies across IDD and other non-IDD populations (McClain et al., 2023), although structural consistency was found for sex, race/ethnicity, and SES (Sparrow et al., 2016). The factor structure literature for the ABAS is even more sparse, with only two published studies, possibly using the same dataset, examining the US, Taiwanese, and Romanian versions and suggesting support for the three major domain factors across versions with evidence of the gender-invariant measurement of the proposed scoring model (Oakland et al., 2013; Wei et al., 2008). Neither the Vineland-3 Social Interaction and Communication domains nor the ABAS-3 Social and Conceptual domains cover a modern conceptualization of core social communication/interaction dimensions identified in the factor-analytic literature (see review above). Using this literature, Figure 1 displays a review of autism and SCI symptom content for the Vineland-3, ABAS-3, and selected autism symptom measures. As can be seen, the Vineland-3 and ABAS-3 show only item (not scale) coverage of the key SCI sub-dimensions.

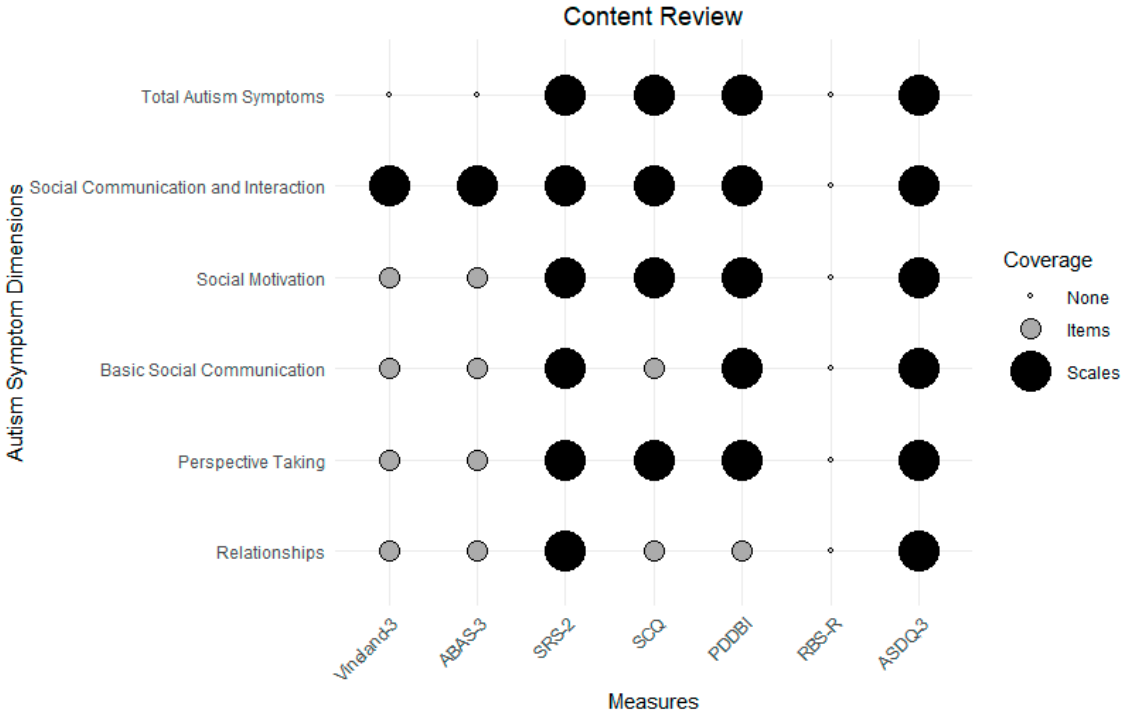


Figure 1. A content review of adaptive function and selected autism symptom measures for total autism symptoms, the SCI domain, and specific SCI dimensions identified in the factor-analytic literature (Chetcuti et al., 2024; Frazier & Hardan, 2017; Phillips et al., 2019; Simmons et al., 2024; Uljarevic et al., 2020a). Note, SRS-2 = Social Responsiveness Scale-Second Edition, SCQ = Social Communication Questionnaire, PDDBI = Pervasive Developmental Disorder Behavior Inventory, RBS-R = Repetitive Behavior Scale-Revised, ASDQ-3 = Autism Symptom Dimensions Questionnaire-Version 3.0.

2.2.3. Reliability

With the caveat of limited evidence for structural validity and its impact on scoring, the broad domain and sub-domain (skill area) scores show very good scale reliability for the Vineland-3 (de Bildt et al., 2005a; Sparrow et al., 2016) and ABAS-3 (Harrison & Oakland, 2015). Conditional reliability via Rasch or item response theory modeling has not been reported in the literature for either measure. Short-term test–retest reproducibility coefficients (12 to 35 days for the Vineland-3; 5 days to 7 weeks for the ABAS-3) are very good for both measures ($r > 0.60$), but neither measure has reported longer-term stability coefficients (3+ months) (American Educational Research Association et al., 2014; Nunnally & Bernstein, 1994). Inter-rater reliability has been found to be good to excellent within forms across interviewers/raters and adequate across different respondents to the same form for the Vineland-3. For the ABAS-3, inter-rater reliability was reported in the manual as good to excellent for raters across forms (in overlapping age bands) as well as good across adult self and informant report versions.

2.2.4. Validity

Both instruments have very good global SCI coverage but, as noted above, the coverage and representation of the more distinct SCI subdomains are limited. For example, the Vineland-3 has many items that assess peer interactions and relationships as well as basic social communication skills (Sparrow et al., 2016). Coverage of social motivation and perspective-taking is also present but less thorough. The ABAS-3 also has very good but not comprehensive coverage. In particular, the conceptual communication subdomain and the social domain of the ABAS-3 have very good coverage of basic social communication skills and relationship skills and some perspective-taking skills but weaker coverage of social motivation (Harrison & Oakland, 2015). Without explicit measurement of this domain, in cases with lower social motivation, initial and ongoing treatment planning could miss the need to focus on increasing affiliative behaviors related to social motivation that are often required to facilitate relationship development. Further, any skill acquisition observed in relationship-building may be challenging to generalize without the child naturally engaging in additional learning opportunities.

Concurrent validity, including the convergent and discriminant validity of the domains and content or skill areas that are measured, is very strong for both instruments (Dupuis et al., 2021), with evidence across multiple studies for the Vineland-3 (de Bildt et al., 2005b; Tillmann et al., 2019). Correlations with autism symptom measures have been observed but have tended to be medium in size and differ across studies and measures (Tillmann et al., 2019; Yang et al., 2016). The Vineland-3 has evidence of domain score sensitivity to change from observational studies and randomized controlled trials (Duncan et al., 2023; Scahill et al., 2016).

2.2.5. Norming

Both instruments have strong normative representation, with samples approximating US population characteristics at the time of collection. Both tools provide age-adjusted standard scores using very large normative samples. The Vineland-3 includes methods that mix a traditional norming approach and accounts for distribution skewness and non-linear age trends. The ABAS-3 uses traditional normative methods, including age bins and matching to the US census. Neither measure uses a modern continuous norming approach (Lenhard & Lenhard, 2021). Continuous norms are a relatively recent innovation with great applicability to developmental assessments. In this approach, norms are built using prediction models (also called regression-based norming) and the models can simultaneously account for multiple demographic factors and their interactions. Our

group has recently shown that a particular form of continuous norming, generalized additive modeling, can result in sensitive normative adjustments with modest neurotypical sample sizes, more accurately accounting for non-linear developmental trajectories and interactions between demographic factors (age by sex) than simpler norming methods (Frazier et al., under review) (Figure S3).

2.2.6. Scoring, Reporting, and Change Measurement

The Vineland-3 provides raw and norm-referenced (age-adjusted) scores for interpretation. The Vineland-3 also provides growth scores for change measurement. Reliable change scores are not available for either instrument. Meaningful change score levels have been recently published for the Vineland-3 (Chatham et al., 2018). Depending on the metric, the anchor-based and distribution-based meaningful change scores tend to be quite small and at or below the standard error of measurement for composite or domain scores. Thus, the very minimal change required for meaningful improvement on the Vineland-3 suggests that most reliable changes are likely to also be viewed as meaningful.

2.3. Pragmatic Considerations

Table 3 presents pragmatic criteria relevant to outcome assessment in real-world CIBI settings, with an overview of the instruments relevant to each characteristic. The Vineland-3 and ABAS-3 offer automated online administration and scoring with the visual display of measure results and interpretive statements to facilitate cross-sectional use. Both measures also provide behavior lists that can inform target identification, although, as noted above, in the content validity section, coverage of relevant behaviors is not as comprehensive as what would be available through a skills-based assessment (Partington, 2010; Sundberg, 2008). Limited intervention strategy guidance is also available, predominantly around the content areas to be addressed. Automated progress monitoring is available in the online system but is limited in terms of tracking longitudinal progress from multiple repeated assessments. The connection of target behavior identification with intervention planning and medical necessity processes is not available. Neither measure offers the ability to aggregate results with other measures that are important for a thorough CIBI outcome assessment to facilitate the clinical workflow.

Table 3. Pragmatic criteria for evaluation of Vineland-3 and ABAS-3 use in real-world CIBI clinical settings.

		Vineland-3	ABAS-3
1	Automated online administration	Yes	Yes
2	Automated scoring	Yes	Yes
3	Visual display of measure results	Yes	Yes
4	Automated interpretative statements	Yes	Yes
5	Automated connection of results with intervention target identification	Partial—item/target behavior lists	Partial—item/target behavior lists
6	Automated clinical guidance for additional assessment, intervention strategy, and referral	Limited intervention strategy guidance	Limited intervention strategy guidance
7	Automated progress monitoring	Limited	Limited
8	Connection of target identification with intervention planning	Not present	Not present
9	Aggregation of other measures and reporting to facilitate clinical workflow	Not present	Not present

Note. CIBI = Comprehensive Intensive Behavioral Intervention. Vineland-3 = Vineland Adaptive Behavior Scales-Third Edition. ABAS-3 = Adaptive Behavior Assessment System-Third Edition.

3. Discussion

3.1. Vineland-3 and ABAS-3 Strengths

The Vineland-3 and ABAS-3 are standardized, norm-referenced assessments with a long history of clinical and research use relevant to autism and other developmental disabilities. The instrument scores have strong scale reliability, test–retest reproducibility, and inter-rater reliability across forms and raters. They also have strong evidence of construct validity for the broad SCI domain and good potential to sensitively detect change (particularly for the Vineland-3). Age-adjusted and growth scale value scores (Vineland-3 only) necessary for intervention applications are available along with basic administration and scoring automation and clinical interpretive guidance. This makes them good choices for cross-sectional assessments, particularly for outpatient mental health settings where the assessment of adaptive behavior for intellectual disability determination is a common clinical question. The Vineland-3 and ABAS-3 also may have value for intervention monitoring in situations where longer SCI assessments are feasible and only measurement of the broad SCI domain is required (Kamphaus & Walden, 2020), with Vineland-3 domain scores having demonstrated evidence of sensitivity to change.

3.2. Vineland-3 and ABAS-3 Weaknesses for the CIBI Context

Relative to their long history of use and strong psychometric properties for cross-sectional use in outpatient mental health contexts, the Vineland-3 and ABAS-3 were not designed for a periodic CIBI outcome assessment process, where measurement of specific SCI dimensions is needed. While other standardized, norm-referenced assessments are sometimes co-administered to compensate for these deficiencies, this adds a significant burden to the process. The Vineland-3 and ABAS-3 also do not include many of the modern measure development procedures relevant to building assessments for a specific purpose or population, such as parent/caregiver and patient/client stakeholder inclusion throughout the process. Unfortunately, in spite of extensive use over many years, the instruments have undergone only partial psychometric evaluations. More specifically, both instruments have received very limited evaluation in terms of their item-level factor structure. When the structure was evaluated using scale scores, originally proposed factor solutions tended to show either poor or adequate rather than excellent fit to support scoring, and importantly, for the Vineland-3, there was also evidence for an alternative structure. In addition, for both instruments, there is only limited evidence of measurement consistency across relevant demographic or clinical subgroups, with the exception of some evidence of subscale level measurement invariance across demographic groups for the Vineland-3. Both measures use Likert frequency scales that are not well-connected to how progress is evaluated in CIBI. Specifically, frequency scaling leaves mid-points of “Sometimes” or “Sometimes” and “Often” that are not linked to the level of prompting or independence that the individual is able to achieve. This makes using item-level scores to inform programmatic decisions unclear. Additionally, content coverage is not sufficient as a stand-alone tool for intervention targeting and goal/objective selection in CIBI and the length of these measures makes supplementation with longer tools less attractive.

The presence of growth value scores (person ability estimates) is a potentially attractive feature of these instruments, but the fact that these scores are only available for broad domain structure makes them less useful when conducting CIBI follow-up evaluations to inform future intervention planning. For example, not having growth scores for specific SCI dimensions makes it difficult to discern if one should continue targeting social motivation and basic social communication skills or move to higher-order cognitive skills. Thus, even though the Vineland-3 domain scores have demonstrated sensitivity to change in several observational and clinical trial contexts, the lack of specific SCI dimension growth

scoring makes it unclear if a broad socialization domain sensitivity to change translates to specific SCI dimensions. Finally, and crucially from the standpoint of usability and implementation in clinical practice, although both instruments have some basic automation and interpretive guidance that facilitates their use in cross-sectional assessment settings, limited functionality is available to increase the efficiency of CIBI progress monitoring. As a result, instrument deployment and the possibility of seamlessly integrating these instruments with other assessments are limited.

3.3. Future Directions

The above-described weaknesses in the most commonly used adaptive functioning measures emphasize the urgent need to rethink how SCI, and more broadly adaptive behavior, is evaluated and considered as part of a larger set of domains/constructs essential for CIBI baseline and follow-up evaluations. Figure 2 presents a list of domains, including SCI, which may be useful to assess for many ASD cases receiving CIBI to improve the treatment process and provide a more in-depth accounting of intervention benefits. The described domains were identified through an iterative process encompassing a systematic review of the literature and input from several stakeholder groups and were part of a broader process aimed at identifying key domains of functioning in ASD and related neurogenetic syndromes, with relevance to appropriate outcome assessment (Joseph et al., 2024). While not all constructs (and associated measurement tools) would be required for all individuals receiving CIBI, it is clear that the current adaptive function measures assess only a subset of relevant SCI behaviors and only a subset of other adaptive behavior constructs that are not aligned with the current understanding of the comprehensive structure and multi-faceted nature of the broad adaptive functioning construct. Simply adding one or more of the existing autism-specific symptom measures leaves many relevant constructs unaccounted for.

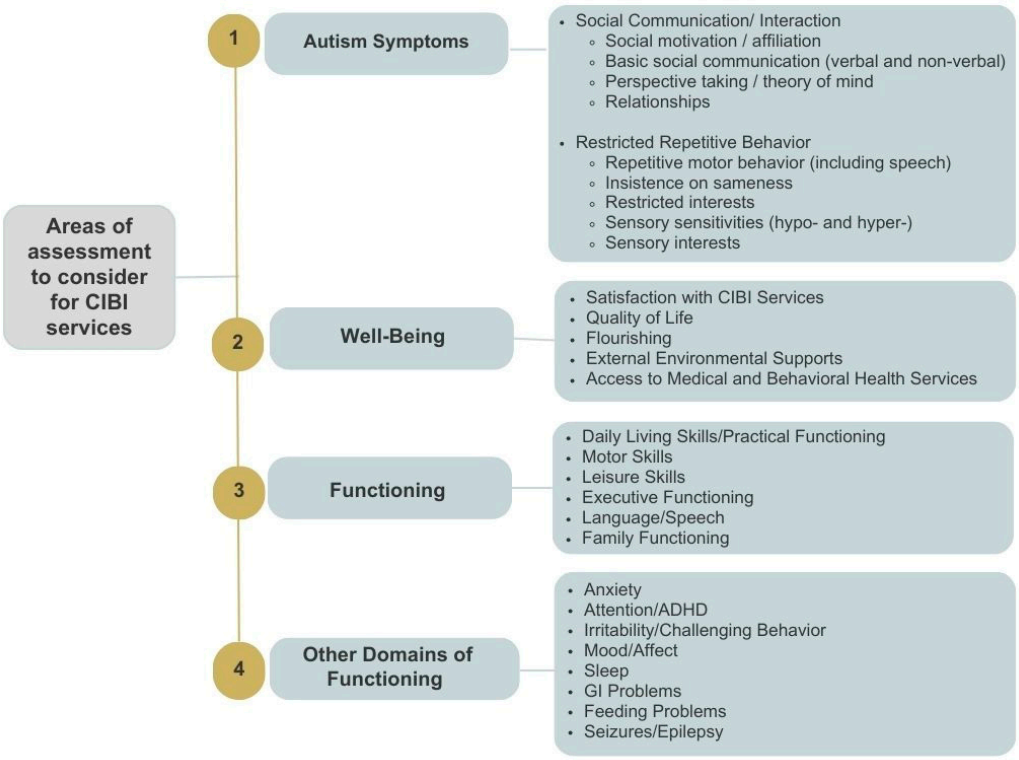


Figure 2. Domains of coverage for initial and periodic CIBI outcome assessment, derived from prior processes for eliciting input from caregiver/patient and clinician/scientist ASD stakeholders. Several

prior large-sample studies support the Social Communication/Interaction framework detailed above (Chetcuti et al., 2024; Frazier & Hardan, 2017; Phillips et al., 2019; Simmons et al., 2024; Uljarevic et al., 2020a). Note, cognition, including general cognitive ability and verbal and nonverbal aspects of intelligence, is a frequent outcome measure in research studies but would be very difficult to include in an initial and periodic CIBI outcome assessment process for everyday practice due to the additional time and expertise required for cognitive test administration and the additional cost added to the process. Unfortunately, as noted above, because both the Vineland-3 and ABAS-3 include a large but not comprehensive list of SCIs, adding skills-based assessments to compensate for incomplete coverage is often not practically feasible. This approach would lead to very long and burdensome assessments without adequately covering all of the relevant domains that are most essential for treatment planning. Importantly, the Vineland-3 and ABAS-3 did not systematically engage with the relevant stakeholders, in particular with autistic individuals, during the initial development phase. In addition to the increased awareness of the need for an inclusive approach to research design, a number of different gold-standard measurement development and validation frameworks, including the PROMIS (Patient-Reported Outcome Measurement Information System) framework, have emphasized the need to engage the relevant stakeholders in order to ensure not only that the measures do not miss any behaviors that are important to capture, but to also facilitate the acceptability of the instrument. Thus, we review three alternative approaches to improving CIBI outcome assessment: combining skills-based assessments with newly developed standardized, norm-referenced assessments; adopting the International Center for Health Outcome Measurement (ICHOM) ASD standard set (Joseph et al., 2024); and developing revised versions of the existing adaptive behavior measures while coupling them with measures of additional domains relevant to CIBI outcome assessment.

3.3.1. Combining Norm-Referenced and Skills-Based Assessments

Many CIBI practitioners, including nearly all BCBAs, are trained in skills-based or criterion-referenced assessment. Although skills and criterion-referenced methods lack norm-referencing and hence the ability to identify intervention targets specifically tied to behavioral domains deviating from normative expectation, they could be used after norm-referenced assessments to identify specific behaviors within behavioral domains that cannot be independently demonstrated by the patient. Thus, a more efficient process would be to deploy a battery of instruments that includes briefer norm-referenced assessments with comprehensive coverage of the relevant domains and a skills-based instrument. This combined norm-referenced and skills-based or criterion-referenced battery would be able to rapidly identify deviations from normative expectation (domains to target) and inform the administration and interpretation of the skills-based assessment, and the overlap between domains identified via norm-referenced assessments and skills-based assessment results could be used to generate plausible intervention target lists. The battery and approach described above leverages one of the strengths of norm-referenced assessments—the ability to identify deviation from developmental expectation (Achenbach, 2001)—and ensures the comprehensive coverage and direct observation of functional behaviors via skills-based or criterion-referenced assessments (Figure S4). For contexts where very detailed assessment is desired, it would also be possible to supplement the above strategy with additional measures and checklists of daily living skills and functional needs (e.g., Essentials for Living, the Early Start Denver Model checklist, etc.).

Identifying a briefer battery of standardized, norm-referenced instruments with good psychometric properties, including sufficient score range and conditional reliability across the domain trait, facilitates repeated deployment of the instruments for the temporally sensitive monitoring of CIBI progress. This broad measurement approach to derive a comprehensive, psychometrically robust, yet brief battery can be effective in tracking the full range of intervention responses, including those occurring in targeted and non-targeted domains. It may seem counter-intuitive to track non-target domains, but CIBI is an intensive and comprehensive strategy and, therefore, is very likely to have "off-target"

effects on a range of behaviors, including attention, anxiety, sleep, the child's and family's quality of life, etc. Thus, it is essential that future practice recommendations and payor policies for CIBI outcome assessment ensure the adequate capture of potentially relevant behavioral domains. The inclusion of a broad monitoring approach with target and non-target domains may also be useful for informing the intervention strategy. For example, even if family and caregiver quality of life is high at baseline evaluation, a future reduction in this domain could signal that parent-mediated intervention is exacerbating stress and potentially reducing the value of the broader treatment strategy. In these scenarios, the clinician may decide to recommend a temporary reduction in parent-mediated goals to reduce stress. Thus, deploying a set of briefer instruments permits more thorough domain coverage and a fuller accounting of CIBI value while also providing information that could signal the need for changes in clinical management. This measurement approach is also likely to provide more value for ongoing clinical quality improvement and value-based care contracting.

Recently, there has been substantial progress in developing and evaluating norm-referenced assessments that assess domains relevant to CIBI outcome assessment, including instruments that are briefer, with strong psychometric properties and the broader coverage of a range of domains and subdomains relevant to CIBI outcome evaluation (Downs et al., 2019; English et al., 2021; Frazier et al., 2023a, 2023b; Gioia et al., 2015; McConachie et al., 2018; Phillips et al., 2019; Reynolds & Kamphaus, 2015; Uljarevic et al., 2022a, 2022b, 2023; Yager & Iarocci, 2013).

Adopting these newly developed instruments, created using best practices and stakeholder inclusion specifically for an autism intervention context, would facilitate the measurement of a wide array of CIBI-relevant neurobehavioral domains. Coupling these measures with an existing skill-based assessment (e.g., VB-MAPP, ABBLS-R, AFLS, PEAK) offers several strengths. First, the inclusion of brief norm-referenced measures of key domains would permit the identification of the patient's level of functioning in each of the SCI dimensions (as well as other relevant domains) and tracking this level over time with less burden on raters than batteries with longer assessments. Second, the inclusion of the skills-based or criterion-referenced assessment permits the direct collection of detailed information on key adaptive behaviors, including SCI-relevant behaviors, to supplement norm-referenced data and facilitate the careful selection of intervention targets. While skills-based assessments have significant psychometric limitations (Padilla et al., 2024), coupling them with brief but psychometrically rigorous norm-referenced assessments covers both major tasks relevant to CIBI outcome assessment—identifying and tracking key behavior domains and choosing intervention targets. Finally, the combination of informant report norm-referenced scales with skills-based or criterion-referenced assessments provides multi-modal assessment (parent/caregiver report, clinician or educator rating, and the direct observation of relevant behaviors). Multi-modal assessment yields a fuller picture of child behavior and provides more data on the therapeutic strategy, setting, and parent/caregiver training that would be difficult to attain with a single modality. The major potential weakness of this approach is that the combination results in a longer and more burdensome assessment process than a single instrument or modality alone. However, this tradeoff is likely acceptable given the need for broad strong coverage of SCI behaviors as well as other adaptive behaviors and child and family functioning.

3.3.2. Adopting the ICHOM Battery

A second option would be to adopt the ICHOM autism spectrum disorder standard set battery (Track A with commercial tools or Track B with a mixture of commercial and free or low-cost tools) (https://www.ichom.org/patient-centered-outcome-measure/

autism-spectrum-disorder/ (accessed on 1 August 2024)) (Joseph et al., 2024). This option is attractive for several reasons. First, the ICHOM battery was developed with stakeholder inclusion and explicit consideration of the domains of assessment that would be most valuable for an autism intervention context, including core autism symptoms, SCI behavior, family functioning, and quality of life. The process also reviewed and prioritized tools based on available psychometric information and used a modified Delphi process to develop agreement on recommendations. The major potential weaknesses of this approach are that (1) it recapitulates criticisms of the Vineland-3 coverage of specific SCI dimensions, as the Social Responsiveness Scale-Second Edition (SRS-2) does not adequately compensate for these limitations, and (2) the lack of inclusion of skills-based or criterion-referenced assessment results in data based largely on a single source (e.g., parent/caregiver) and provides limited data for SCI intervention target selection. Furthermore, many of the instruments included in the ICHOM battery lack the pragmatic considerations described above (e.g., no automated administration, scoring, or reporting).

3.3.3. Revision of Vineland-3 and ABAS-3 for CIBI

A third option would be to revise the existing adaptive behavior instruments for a CIBI outcome assessment context. Given the above-reviewed limited construct coverage of current adaptive behavior instruments, this approach would require augmenting existing item content to ensure the sufficient coverage of SCI subdimensions to support scoring as well as the potential removal of existing items/subscales (e.g., the Vineland-3 written expression subscale). The revised instruments would need to undergo significant psychometric evaluation before implementation, with particular emphasis on developing a replicable factor structure and measurement invariance across key demographic and clinical characteristics. These revisions would require significant time and resources and this option would also need to consider either (i) supplementing revised adaptive function measures with instruments that evaluate other relevant domains, such as the Behavior Assessment System for Children-Third Edition for other psychopathology domains, such as the WHOQoL-BREF (McConachie et al., 2018) or the QI-Disability (Downs et al., 2019) to evaluate quality of life/flourishing, the Behavior Rating Inventory of Executive Function (BRIEF-2) for executive functioning (Gioia et al., 2015), and a measure of core autism symptoms; (ii) or developing additional item sets that would be incorporated into the revised adaptive assessments to assess these additional content areas. Similar to the ICHOM battery, this approach would also focus on informant report without direct collection of key SCI behaviors. Thus, while this option may have strong long-term potential if the necessary revision processes are undertaken, it is not likely to be an adequate short-term solution. Finally, this approach would also require addressing the current practical limitations of the Vineland-3 and ABAS-3, such as longitudinal score monitoring, linking results to potential intervention targets, and aiding the clinician in detecting reliable change to inform clinical management decisions.

4. Conclusions

In contrast to their strong utility in cross-sectional outpatient mental health assessment settings, the Vineland-3 and ABAS-3 present with a range of psychometric and conceptual limitations that significantly impact their utility as outcome assessments. These legacy adaptive function measures have long administration times; lack an in-depth coverage of key aspects of adaptive functioning that are aligned with the current theoretical models and empirical factorial evidence and findings from the studies that have examined the outcome priorities of the autistic community and their families; have inadequate psychometric information to support their scoring, interpretation, and monitoring; and provide only a

subset of the practical implementation features that would facilitate use in everyday CIBI practice. Providers and payors should reconsider the adoption of these measures as part of a CIBI outcome assessment battery. Instead, a more optimized CIBI assessment approach is recommended that collects multi-modal information and leverages the strengths of newly developed norm-referenced assessments for the autism behavioral intervention context coupled with skills-based and neurodiversity-affirming assessments that provide deep coverage of key social and functional domains relevant to choosing specific intervention targets. In addition, once such assessments are developed and stringently validated, it will be essential to explore and establish their cross-cultural validity and, if needed, make appropriate updates and refinements.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/bs15060722/s1, Table S1. Comparison of SCI-relevant items on adaptive function and autism symptom measures. Table S2. Specific SCI dimensions identified through factor analytic studies. Figure S1. Flow diagram for the initial referral and CIBI outcome assessment process. Figure S2. Visual depiction of the application of initial and periodic CIBI outcome assessments to facilitate treatment planning and ongoing clinical management. Figure S3. Example continuous norming for total sleep problems using generalized additive modeling with an interaction between age and sex across a wide age range (2 to 90) in a medium-sized normative sample (N=536). Figure S4. Proposed process for integrating results from briefer standardized, norm-referenced assessments, including measures of adaptive function, with results from skills-based / criterion-referenced assessments to develop a CIBI treatment protocol.

Author Contributions: Conceptualization, T.W.F., E.A.Y. and M.U.; methodology, T.W.F. and M.U.; resources, T.W.F.; writing—original draft preparation, T.W.F. and M.U.; writing—review and editing, all authors; visualization, T.W.F. and M.U.; supervision, T.W.F.; project administration, T.W.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Acknowledgments: The authors would like to acknowledge Rebecca A. Womack for her contribution to the early drafts of the manuscript. We would also like to thank the caregivers and primary participants for their time and effort in completing the study.

Conflicts of Interest: Author Allison R. Frazier was employed by the company Elevate Learning. Frazier has received funding or research support from, acted as a consultant to, received travel support from, and/or received a speaker's honorarium from the PTEN Research Foundation, SYNGAP Research Fund, Malan Syndrome Foundation, ADNP Kids Research Foundation, Quadrant Biosciences, Autism Speaks, Impel NeuroPharma, F. Hoffmann-La Roche AG Pharmaceuticals, the Cole Family Research Fund, Simons Foundation, Ingalls Foundation, Forest Laboratories, Ecoeos, IntegraGen, Kugona LLC, Shire Development, Bristol-Myers Squibb, Roche Pharma, MaraBio, Scioto Biosciences, National Institutes of Health, and the Brain and Behavior Research Foundation; has equity options in Quadrant Biosciences/Autism Analytica, MaraBio, and Springtide; is co-founder and has equity interest in AI-Measures; and has an investor stake in Autism EYES and iSCAN-R. Youngstrom is the co-founder and executive director of Helping Give Away Psychological Science, a 501c3; he has consulted on psychological assessment with Signant Health, and received royalties from the American Psychological Association and Guilford Press; and he holds equity in AI-Measures and Joe Startup Technologies. Frazier and Womack have equity interest in AI-Measures. Uljarevic has equity interest in iSCAN-R and AI-Measures. The above-listed organizations had no role in the design of the study; in the collection, analyses, or interpretation of the data; in the writing of the manuscript; or in the decision to publish the results.

References

Achenbach, T. M. (2001). What are norms and why do we need valid ones? *Clinical Psychology: Science & Practice*, *8*(4), 446–450. [CrossRef]

American Educational Research Association, American Psychological Association & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association, American Psychological Association, & National Council on Measurement in Education.

Baiden, K. M. P., Williams, Z. J., Schuck, R. K., Dwyer, P., & Wang, M. (2024). The social validity of behavioral interventions: Seeking input from autistic adults. *Journal of Autism and Developmental Disorders*, *55*, 1172–1186. [CrossRef] [PubMed]

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quinonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health*, *6*, 149. [CrossRef]

Chatham, C. H., Taylor, K. I., Charman, T., Liogier D'ardhuy, X., Eule, E., Fedele, A., Hardan, A. Y., Loth, E., Murtagh, L., Del Valle Rubido, M., San Jose Caceres, A., Sevigny, J., Sikich, L., Snyder, L., Tillmann, J. E., Ventola, P. E., Walton-Bowen, K. L., Wang, P. P., Willgoss, T., & Bolognani, F. (2018). Adaptive behavior in autism: Minimal clinically important differences on the Vineland-II. *Autism Research*, *11*(2), 270–283. [CrossRef]

Chetcuti, L., Hardan, A. Y., Spackman, E., Frazier, T. W., Youngstrom, E. A., & Uljarevic, M. (2024). Characterizing transdiagnostic processes underlying the drive to socially engage: Multimeasurement factor analytic investigation. *Journal of Child and Adolescent Psychiatry Open*. [CrossRef]

Clausen, C. E., Leventhal, B. L., Nytro, O., Koposov, R., Westbye, O. S., Rost, T. B., Bakken, V., Koochakpour, K., Thorvik, K., & Skokauskas, N. (2021). Clinical Decision Support Systems: An Innovative Approach to Enhancing Child and Adolescent Mental Health Services. *Journal of the American Academy of Child and Adolescent Psychiatry*, *60*(5), 562–565. [CrossRef] [PubMed]

Cooper, J., Heron, T., & Heward, W. (2007). *Applied behavior analysis* (2nd ed.). Pearson/Merrill-Prentice Hall.

Council of Autism Service Providers. (2024). *Applied behavior analysis practice guidelines for the treatment of autism spectrum disorder: Guidelines for healthcare funders, regulatory bodies, service providers, and consumers* (3rd ed.). Available online: https://www.casproviders.org/asd-guidelines (accessed on 1 August 2024).

de Bildt, A., Kraijer, D., Sytema, S., & Minderaa, R. (2005a). The psychometric properties of the Vineland Adaptive Behavior Scales in children and adolescents with mental retardation. *Journal of Autism and Developmental Disorders*, *35*(1), 53–62. [CrossRef]

de Bildt, A., Sytema, S., Kraijer, D., Sparrow, S., & Minderaa, R. (2005b). Adaptive functioning and behaviour problems in relation to level of education in children and adolescents with intellectual disability. *Journal of Intellectual Disability Research*, *49 Pt 9*, 672–681. [CrossRef]

Downs, J., Jacoby, P., Leonard, H., Epstein, A., Murphy, N., Davis, E., Reddihough, D., Whitehouse, A., & Williams, K. (2019). Psychometric properties of the Quality of Life Inventory-Disability (QI-Disability) measure. *Quality of Life Research*, *28*(3), 783–794. [CrossRef]

Duncan, A., Meinzen-Derr, J., Ruble, L., Fassler, C., & Stark, L. J. (2023). A randomized clinical trial targeting daily living skills in autistic adolescents without an intellectual disability before the transition to adulthood. *Journal of Developmental and Behavioral Pediatrics*, *44*(9), e590–e596. [CrossRef]

Dupuis, A., Moon, M. J., Brian, J., Georgiades, S., Levy, T., Anagnostou, E., Nicolson, R., Schachar, R., & Crosbie, J. (2021). Concurrent validity of the ABAS-II questionnaire with the Vineland II interview for adaptive behavior in a pediatric ASD sample: High correspondence despite systematically lower scores. *Journal of Autism and Developmental Disorders*, *51*(5), 1417–1427. [CrossRef] [PubMed]

English, M. C. W., Gignac, G. E., Visser, T. A. W., Whitehouse, A. J. O., Enns, J. T., & Maybery, M. T. (2021). The Comprehensive Autistic Trait Inventory (CATI): Development and validation of a new measure of autistic traits in the general population. *Molecular Autism*, *12*(1), 37. [CrossRef] [PubMed]

Farmer, C. A., Kaat, A. J., Thurm, A., Anselm, I., Akshoomoff, N., Bennett, A., Berry, L., Bruchey, A., Barshop, B. A., Berry-Kravis, E., Bianconi, S., Cecil, K. M., Davis, R. J., Ficicioglu, C., Porter, F. D., Wainer, A., Goin-Kochel, R. P., Leonczyk, C., Guthrie, W., … Miller, J. S. (2020). Person ability scores as an alternative to norm-referenced scores as outcome measures in studies of neurodevelopmental disorders. *American Journal on Intellectual and Developmental Disabilities*, *125*(6), 475–480. [CrossRef]

Farmer, R. L., Floyd, R. G., & McNicholas, P. J. (2021). Is the Vineland-3 comprehensive interview form a multidimensional or unidimensional scale? Structural analysis of subdomain scores across early childhood to adulthood. *Assessment*, *28*(7), 1848–1864. [CrossRef] [PubMed]

FDA. (2009). *Patient-reported outcome measures: Use in medical product development to support labeling claims*. United States Food and Drug Administration.

Forbes, G., Kent, R., Charman, T., Baird, G., Pickles, A., & Simonoff, E. (2023). How do autistic people fare in adult life and can we predict it from childhood? *Autism Research*, *16*(2), 458–473. [CrossRef]

Frazier, T. W., Busch, R. M., Klaas, P., Lachlan, K., Jeste, S., Kolevzon, A., Loth, E., Harris, J., Speer, L., Pepper, T., Anthony, K., Graglia, J. M., Delagrammatikas, C., Bedrosian-Sermone, S., Beekhuyzen, J., Smith-Hicks, C., Sahin, M., Eng, C., Hardan, A. Y., & Uljarevic, M. (2023a). Development of informant-report neurobehavioral survey scales for PTEN hamartoma tumor syndrome and related neurodevelopmental genetic syndromes. *The American Journal of Medical Genetics Part A*, *191*(7), 1741–1757. [CrossRef]

Frazier, T. W., Dimitropoulos, A., Abbeduto, L., Armstrong-Brine, M., Kralovic, S., Shih, A., Hardan, A. Y., Youngstrom, E. A., Uljarevic, M., & Quadrant Biosciences—As You Are Team. (2023b). The autism symptom dimensions questionnaire: Development and psychometric evaluation of a new, open-source measure of autism symptomatology. *Developmental Medicine and Child Neurology*, *65*(8), 1081–1092. [CrossRef]

Frazier, T. W., & Hardan, A. Y. (2017). Equivalence of symptom dimensions in females and males with autism. *Autism*, *21*(6), 749–759. [CrossRef]

Frazier, T. W., Huba, K., Frazier, A. R., Womack, R. A., Youngstrom, E. A., Chetcuti, L. C., Hardan, A. Y., & Uljarevic, M. (under review). Maximizing accurate detection of divergence from normative expectation in behavioral intervention outcome assessment.

Frazier, T. W., Ratliff, K. R., Gruber, C., Zhang, Y., Law, P. A., & Constantino, J. N. (2014). Confirmatory factor analytic structure and measurement invariance of quantitative autistic traits measured by the Social Responsiveness Scale-2. *Autism*, *18*(1), 31–44. [CrossRef]

Frazier, T. W., Youngstrom, E. A., Kubu, C. S., Sinclair, L., & Rezai, A. (2008). Exploratory and confirmatory factor analysis of the autism diagnostic interview-revised. *Journal of Autism and Developmental Disorders*, *38*(3), 474–480. [CrossRef]

Frazier, T. W., Youngstrom, E. A., Speer, L., Embacher, R., Law, P., Constantino, J., Findling, R. L., Hardan, A. Y., & Eng, C. (2012). Validation of proposed DSM-5 criteria for autism spectrum disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, *51*(1), 28–40.e23. [CrossRef] [PubMed]

Georgiades, S., Szatmari, P., Boyle, M., Hanna, S., Duku, E., Zwaigenbaum, L., Bryson, S., Fombonne, E., Volden, J., Mirenda, P., Smith, I., Roberts, W., Vaillancourt, T., Waddell, C., Bennett, T., & Thompson, A. (2012). Investigating phenotypic heterogeneity in children with autism spectrum disorder: A factor mixture modeling approach. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *54*(2), 206–215. [CrossRef]

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2015). *Behavior rating inventory of executive function–second edition (BRIEF-2)*. Psychological Assessment Resources.

Happe, F., Cook, J., & Bird, G. (2017). The structure of social cognition: In(ter)dependence of sociocognitive processes. *Annual Review of Psychology*, *68*, 243–267. [CrossRef]

Harrison, P., & Oakland, T. (2015). *Adaptive behavior assessment system* (3rd ed.). Western Psychological Services.

Huber, L., Plotner, M., & Schmitz, J. (2019). Social competence and psychopathology in early childhood: A systematic review. *European Child and Adolescent Psychiatry*, *28*(4), 443–459. [CrossRef]

Hume, K., Steinbrenner, J. R., Odom, S. L., Morin, K. L., Nowell, S. W., Tomaszewski, B., Szendrey, S., McIntyre, N. S., Yucesoy-Ozkan, S., & Savage, M. N. (2021). Evidence-based practices for children, youth, and young adults with autism: Third generation review. *Journal of Autism and Developmental Disorders*, *51*(11), 4013–4032. [CrossRef]

Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., Sanislow, C., & Wang, P. (2010). Research domain criteria (RDoC): Toward a new classification framework for research on mental disorders. *American Journal of Psychiatry*, *167*(7), 748–751. [CrossRef] [PubMed]

Joseph, A., Chong, I., Das-Gupta, Z., Bandeira de Lima, C., Dixon, D., Dovbnya, S., Fittro, E., Gerhardt, P., Huang, W., Josephson, B., Li, D., Martin, N., Mukerji, S., Rodriguez, K., Rue, H., Strunk, K., Tarbox, J., Vadgama, Y., Valentino, A., . . . Willis, S. (2024). Development of a standardized set of outcomes for autism spectrum disorder: The International Consortium for Health Outcomes Measurement (ICHOM). *Research in Autism Spectrum Disroders*, *117*, 102451. [CrossRef]

Kamphaus, R., & Walden, E. (2020). Intellectual disability. In *Assessment of disorders in childhood and adolescence* (5th ed., pp. 358–373). Guilford Press.

Lenhard, W., & Lenhard, A. (2021). Improvement of norm score quality via regression-based continuous norming. *Educational and Psychological Measurement*, *81*(2), 229–261. [CrossRef] [PubMed]

Lord, C., Charman, T., Havdahl, A., Carbone, P., Anagnostou, E., Boyd, B., Carr, T., de Vries, P. J., Dissanayake, C., Divan, G., Freitag, C. M., Gotelli, M. M., Kasari, C., Knapp, M., Mundy, P., Plank, A., Scahill, L., Servili, C., Shattuck, P., . . . McCauley, J. B. (2021). The lancet commission on the future of care and clinical research in autism. *Lancet*, *399*(10321), 271–334. [CrossRef]

Mandy, W. P., Charman, T., & Skuse, D. H. (2012). Testing the construct validity of proposed criteria for DSM-5 autism spectrum disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, *51*(1), 41–50. [CrossRef]

Marsh, H. W., Hau, K. T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indices and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling*, *11*, 320–341. [CrossRef]

Marsh, H. W., Morin, A. J., Parker, P. D., & Kaur, G. (2014). Exploratory structural equation modeling: An integration of the best features of exploratory and confirmatory factor analysis. *The Annual Review of Clinical Psychology*, *10*, 85–110. [CrossRef] [PubMed]

McClain, M. B., Schwartz, S. E., Bera, J., Farmer, R. L., Serang, S., Harris, B., & Golson, M. E. (2023). Vineland-3 measurement non-invariance in children with and without intellectual and developmental disabilities. *American Journal on Intellectual and Developmental Disabilities*, *128*(4), 334–343. [CrossRef]

McConachie, H., Mason, D., Parr, J. R., Garland, D., Wilson, C., & Rodgers, J. (2018). Enhancing the validity of a quality of life measure for autistic people. *Journal of Autism and Developmental Disorders*, *48*(5), 1596–1611. [CrossRef] [PubMed]

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). McGraw-Hill, Inc.

Oakland, T., Illiescu, D., Chen, H.-Y., & Chen, J. H. (2013). Cross-national assessment of adaptive behavior in three countries. *Journal of Psychoeducational Assessment*, *31*(5), 435–447. [CrossRef]

Padilla, K. L., Sarno, J., & Kazemi, E. (2024). Assessment use in applied behvior analysis. In L. M. Toby, & E. S. Ranade (Eds.), *Psychology essentials for behvior analysts*. Routledge. [CrossRef]

Pandolfi, V., & Magyar, C. I. (2021). Vineland-3 structural validity and interpretability of domain scores: Implications for practitioners assessing adolescents with developmental conditions. *American Journal on Intellectual and Developmental Disabilities*, *126*(3), 216–229. [CrossRef]

Partington, J. W. (2010). *The ABLLS-R—The assessment of basic language and learning skills-revised*. Behavior Analysts, Inc.

Phillips, J. M., Uljarevic, M., Schuck, R. K., Schapp, S., Solomon, E. M., Salzman, E., Allerhand, L., Libove, R. A., Frazier, T. W., & Hardan, A. Y. (2019). Development of the stanford social dimensions scale: Initial validation in autism spectrum disorder and in neurotypicals. *Molecular Autism*, *10*, 48. [CrossRef]

Rettew, D. C., Lynch, A. D., Achenbach, T. M., Dumenci, L., & Ivanova, M. Y. (2009). Meta-analyses of agreement between diagnoses made from clinical evaluations and standardized diagnostic interviews. *International Journal of Methods in Psychiatric Research*, *18*(3), 169–184. [CrossRef]

Reynolds, C. R., & Kamphaus, R. W. (2015). *BASC-3: Behavior assessment system for children* (3rd ed.). NCS Pearson, Inc.

Ridout, S., & Eldevik, S. (2023). Measures used to assess treatment outcomes in children with autism receiving early and intensive behavioral interventions: A review. *Review Journal of Autism and Developmental Disorders*, *11*, 607–619. [CrossRef]

Rogers, S. J., Yoder, P., Estes, A., Warren, Z., McEachin, J., Munson, J., Rocha, M., Greenson, J., Wallace, L., Gardner, E., Dawson, G., Sugar, C. A., Hellemann, G., & Whelan, F. (2021). A multisite randomized controlled trial comparing the effects of intervention intensity and intervention style on outcomes for young children with autism. *Journal of the American Academy of Child and Adolescent Psychiatry*, *60*(6), 710–722. [CrossRef]

Scahill, L., Bearss, K., Lecavalier, L., Smith, T., Swiezy, N., Aman, M. G., Sukhodolsky, D. G., McCracken, C., Minshawi, N., Turner, K., Levato, L., Saulnier, C., Dziura, J., & Johnson, C. (2016). Effect of parent training on adaptive behavior in children with autism spectrum disorder and disruptive behavior: Results of a randomized trial. *Journal of the American Academy of Child and Adolescent Psychiatry*, *55*(7), 602–609.e603. [CrossRef]

Schreibman, L., Dawson, G., Stahmer, A. C., Landa, R., Rogers, S. J., McGee, G. G., Kasari, C., Ingersoll, B., Kaiser, A. P., Bruinsma, Y., McNerney, E., Wetherby, A., & Halladay, A. (2015). Naturalistic developmental behavioral interventions: Empirically validated treatments for autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *45*(8), 2411–2428. [CrossRef]

Schuck, R. K., Baiden, K. M. P., Wang, A. T., Olis, S., Ingram, C., & Fisher, G. (2024a). Assessment of adverse events, side effects, and social validity in evidence-based behavioral interventions for autistic students. *Review of Research in Education*, *48*(1), 154–190. [CrossRef]

Schuck, R. K., Choi, S., Baiden, K. M., Dwyer, P., & Uljarevic, M. (2024b). The neurodiversity attitudes questionnaire: Development and initial validation. *Autism*, *28*(11), 2821–2833. [CrossRef] [PubMed]

Schuck, R. K., Tagavi, D. M., Baiden, K. M. P., Dwyer, P., Williams, Z. J., Osuna, A., Ferguson, E. F., Jimenez Munoz, M., Poyser, S. K., Johnson, J. F., & Vernon, T. W. (2022). Neurodiversity and autism intervention: Reconciling perspectives through a naturalistic developmental behavioral intervention framework. *Journal of Autism and Developmental Disorders*, *52*(10), 4625–4645. [CrossRef]

Shuster, J., Perry, A., Bebko, J., & Toplak, M. E. (2014). Review of factor analytic studies examining symptoms of autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *44*(1), 90–110. [CrossRef]

Simmons, G. L., Corbett, B. A., Lerner, M. D., Wofford, K., & White, S. W. (2024). Social competence in autism: A structural equation modeling approach. *Autism Research*, *17*(4), 761–774. [CrossRef] [PubMed]

Smith, T., Klorman, R., & Mruzek, D. W. (2015). Predicting outcome of community-based early intensive behavioral intervention for children with autism. *Journal of Abnormal Child Psychology*, *43*(7), 1271–1282. [CrossRef]

Snow, A. V., Lecavalier, L., & Houts, C. (2009). The structure of the autism diagnostic interview-revised: Diagnostic and phenotypic implications. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *50*(6), 734–742. [CrossRef] [PubMed]

Sorensen, O., Fjell, A. M., & Walhovd, K. B. (2023). Longitudinal modeling of age-dependent latent traits with generalized additive latent and mixed models. *Psychometrika*, *88*(2), 456–486. [CrossRef] [PubMed]

Sparrow, S. S., Cicchetti, D. V., & Saulnier, C. A. (2016). *Vineland adaptive behavior scales, third edition (Vineland-3)*. Pearson.

Sulek, R., Edwards, C., Monk, R., Patrick, L., Pillar, S., & Waddington, H. (2024). Community priorities for outcomes targeted during professional supports for autistic children and their families. *Journal of Autism and Developmental Disorders*, *55*, 1890–1901. [CrossRef] [PubMed]

Sundberg, M. L. (2008). *Verbal behavior milestones assessment and placement program: The VB-MAPP* (2nd ed.). AVB Press.

Tillmann, J., San Jose Caceres, A., Chatham, C. H., Crawley, D., Holt, R., Oakley, B., Banaschewski, T., Baron-Cohen, S., Bolte, S., Buitelaar, J. K., Durston, S., Ham, L., Loth, E., Simonoff, E., Spooren, W., Murphy, D. G., Charman, T., & EU-AIMS LEAP group. (2019). Investigating the factors underlying adaptive functioning in autism in the EU-AIMS longitudinal european autism project. *Autism Research*, *12*(4), 645–657. [CrossRef] [PubMed]

Trembath, D., Varcin, K., Waddington, H., Sulek, R., Bent, C., Ashburner, J., Eapen, V., Goodall, E., Hudry, K., Roberts, J., Silove, N., & Whitehouse, A. (2023). Non-pharmacological interventions for autistic children: An umbrella review. *Autism*, *27*(2), 275–295. [CrossRef] [PubMed]

Uljarevic, M., Cai, R. Y., Hardan, A. Y., & Frazier, T. W. (2022a). Development and validation of the executive functioning scale. *Front Psychiatry*, *13*, 1078211. [CrossRef]

Uljarevic, M., Frazier, T. W., Jo, B., Scahill, L., Youngstrom, E. A., Spackman, E., Phillips, J. M., Billingham, W., & Hardan, A. (2023). Dimensional assessment of restricted and repetitive behaviors: Development and preliminary validation of a new measure. *Journal of the American Academy of Child and Adolescent Psychiatry*, *62*(5), 568–581. [CrossRef]

Uljarevic, M., Frazier, T. W., Phillips, J. M., Jo, B., Littlefield, S., & Hardan, A. Y. (2020a). Mapping the research domain criteria social processes constructs to the social responsiveness scale. *Journal of the American Academy of Child and Adolescent Psychiatry*, *59*(11), 1252–1263.e1253. [CrossRef]

Uljarevic, M., Frazier, T. W., Phillips, J. M., Jo, B., Littlefield, S., & Hardan, A. Y. (2021). Quantifying research domain criteria social communication subconstructs using the social communication questionnaire in youth. *Journal of Clinical Child & Adolescent Psychology*, *50*(5), 609–618. [CrossRef]

Uljarevic, M., Phillips, J. M., Schuck, R. K., Schapp, S., Solomon, E. M., Salzman, E., Allerhand, L., Libove, R. A., Frazier, T. W., & Hardan, A. Y. (2020b). Exploring social subtypes in autism spectrum disorder: A preliminary study. *Autism Research*, *13*(8), 1335–1342. [CrossRef]

Uljarevic, M., Spackman, E. K., Cai, R. Y., Paszek, K. J., Hardan, A. Y., & Frazier, T. W. (2022b). Daily living skills scale: Development and preliminary validation of a new, open-source assessment of daily living skills. *Front Psychiatry*, *13*, 1108471. [CrossRef]

Waddington, H., Minnell, H., Patrick, L., van Der Meer, L., Monk, R., Woods, L., & Whitehouse, A. J. (2024). Community perspectives on the appropriateness and importance of support goals for young autistic children. *Autism*, *28*(2), 316–326. [CrossRef]

Wei, Y., Oakland, T., & Algina, J. (2008). Multigroup confirmatory factor analysis for the adaptive behavior assessment system-II parent form, ages 5–21. *American Journal of Mental Retardation*, *113*(3), 178–186. [CrossRef] [PubMed]

Wilkinson, E., Farmer, C., Kleiman, E., & Bal, V. H. (2023). Factor structure of the VABS-3 comprehensive parent/caregiver form in autistic individuals: Poor fit of three-factor and unidimensional models. *Autism*, *28*(3), 13623613231179288. [CrossRef]

Yager, J., & Iarocci, G. (2013). The development of the multidimensional social competence scale: A standardized measure of social competence in autism spectrum disorders. *Autism Research*, *6*(6), 631–641. [CrossRef] [PubMed]

Yang, S., Paynter, J. M., & Gilmore, L. (2016). Vineland adaptive behavior scales: II profile of young children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *46*(1), 64–73. [CrossRef] [PubMed]

Youngstrom, E. A., & Prinstein, M. J. (2020). Introduction to evidence-based assessment: A recipe for success. In M. J. P. E. A. Youngstrom, E. J. Mash, & R. Barkley (Eds.), *Assessment of disorders in childhood and adolescence* (5th ed., pp. 3–29). Guilford Press.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.