EXPLORATORY AND CONFIRMATORY FACTOR ANALYSIS OF THE REPETITIVE BEHAVIOR SCALE – REVISED (RBS-R) IN AN AUTISM SPECTRUM DISORDER SAMPLE WITH RATINGS COMPLETED BY SPECIAL EDUCATION STAFF

By

Megan Michelle Stoll

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

School Psychology – Doctor of Philosophy

2023

ABSTRACT

Autism spectrum disorder (ASD) is characterized in the Diagnostic and Statistical Manual of Mental Disorders (DSM-5; APA, 2013) by deficits in social communication and the presence of restricted and repetitive behaviors and interests (RRBIs). Despite the mandatory requirement of both categories of impairment for a diagnosis of ASD, research regarding RRBIs is minimal, when compared to that of social impairments. The Repetitive Behavior Scale-Revised (RBS-R; Bodfish et al., 2000) is a rating scale designed to capture the broad range and scope of RRBIsin individuals with ASD. Among the few RRBI measures available, researchers have used the RBS-R as an outcome measure in several intervention studies. Despite its relative popularity, RBS-R validity research has yet to yield a consensus regarding the tool's factor structure. Such consensus is needed to clarify the number of subscales and their constructs, as well as establish item-to-subscale assignments. Previous factor analytic studies (Bishop et al., 2013; Bodfish et al., 2000; Lam & Aman, 2007; Mirenda et al., 2010; Russell et al., 2019; Hooker et al., 2019; Sturm et al., 2022), have used caregiver ratings on the RBS-R in ASD samples, yielding a variety of factor solutions (i.e., four-, five-, or six-factor solutions). Though variable, analyses have produced more support for a five-factor solution among caregiver ratings of repetitive behavior (Bishop et al., 2013; Hooker et al. 2019; Lam & Aman, 2007; Mirenda et al., 2010; Sturm et al., 2022). Additional factor analytic studies conducted using non-English versions of the RBS-R in other cultural contexts (Georgiades et al., 2010; He et al., 2019; Kästel et al., 2020), using caregiver ratings of ASD samples also yielded inconsistencies in the number of factors present (i.e., two, four, and six factors). Only one published factor analytic study involved the use of school staff ratings of the RBS-R in an ASD sample. This study (Martínez-González & Piqueras, 2017) utilized the Spanish version of the instrument in Spain and authors

retained a six-factor solution. This has left questions about the range of factor solutions examined across studies, as well as possible differences in factor structure due to rater type. The lack of consensus, in addition to the lack of school raters, warrants independent factor analyses to determine the most appropriate, valid factor structure of the RBS-R among school staff raters. The present dissertation involves two independent studies. Study one utilized an exploratory factor analysis (EFA) of the RBS-R items with an ASD sample (*N* = 234), rated by special education staff. The EFA resulted in a five-factor solution that demonstrated some similarities to existing factor models, though presented clear differences not previously observed in the literature. Study two involved applying confirmatory factor analysis (CFA) in a second independent ASD validation sample (*N* = 233) from the same agency used in study one. CFA procedures were used to test the fit of the RBS-R factor solution retained from study one across five fit indices and to compare the fit of that model to others available in the literature. CFA results suggested that the five-factor model from study one adequately fit the sample data. AIC and BIC indices indicated that the study one model was the best-fitting model compared to the existing models. Findings underscore the possibility that the current structure of the RBS-R may not be most viable, when using school staff ratings. Findings also suggest the presence of a new, unique factor resulting from dividing items from the original RBS-R self-injurious behavior factor. Further, findings leave room for future studies to continue to consider the factor structure and fit of models of the RBS-R based on school staff ratings in samples of those with ASD.

ACKNOWLEDGEMENTS

I would like to thank my family for their love, support, and understanding of my lifelong role as a student. Mom – words cannot express how lucky I am to have you as my lifelong support system and sounding board. To my husband, Jason – thank you for being my source of calmness, patience, and understanding. I could have done it without you, but I would have never wanted to. Thank you to my friends Kim, Savannah, Adam, Lizzy, Andrew, Tessa, Nicole, Emma, Kristina, and the 15 Alumni. Thank you to my cohort mates and friends, Sarina and Tyler, for perfectly balancing out my worries and productivity throughout the last six years.

Thank you to my dissertation committee and MSU School Psychology faculty. To my advisor, Dr. Martin Volker – thank you for your expertise, kindness, and mentorship. I am very thankful for your guidance and investment in my professional development. To Dr. Kristin Rispoli, Dr. Gloria Lee, and Dr. Connie Sung – thank you for your understanding, encouragement, and thoughtful feedback throughout the dissertation process.

Thank you to my undergraduate advisors for inspiring me to pursue my professional career. Dr. Jennifer Lodi-Smith, Dr. Jonathan Rodgers, and Dr. Charles Goodsell – your influence on my research interests is unmeasurable.

To my research team: Dr. Nicole Bergamo Isbell, Dr. Janelle Taylor Youngdahl, and Shelby Brennan – thank you for your knowledge, guidance, and assistance with my projects.

TABLE OF CONTENTS

CHAPTER ONE: INTRODUCTION

Autism spectrum disorder (ASD) is designated as a neurodevelopmental disorder within the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* (DSM-5; American Psychiatric Association; APA, 2013). According to the Centers for Disease Control and Prevention (CDC), the prevalence of ASD, which is estimated currently at 1 in 54 among eight-year-old children, has been steadily increasing over the last two decades (Baio et al., 2018; Maenner et al., 2020). Core diagnostic features of the disorder consist of significant impairments in social communication and the presence of restricted, repetitive, or stereotyped patterns of behaviors, interests, or activities (APA, 2013). Social communication and interaction deficits may present as difficulties in social-emotional reciprocity, nonverbal communicative behaviors (used for social interaction), and developing and maintaining relationships. In the domain of restricted, repetitive patterns of behaviors and interests (RRBI), common presentations include repetitive motor movements or use of objects, insistence on sameness, ritualized patterns of verbal and nonverbal behavior, highly restricted and fixated interests, and hyper- or hypo-reactivity to sensory input (APA, 2013).

ASD Assessment Practices

With the rising prevalence rates in the United States, assessment tools used for ASD evaluations have faced periodic modifications to enhance the quality of diagnostic decision-making for individuals with ASD. The field requires a more uniform assessment process that is valid and efficient to augment the quality of ASD assessment (Matson & Kozlowski, 2011). At present, several tools assess a range of features and skills in individuals with ASD who may present with or without co-morbid intellectual disability (ID). For diagnostic and initial intervention purposes, broad assessments typically include an evaluation of core features, medical and developmental history, current cognitive ability, adaptive behavior, speech/language and hearing, and motor skills (APA, 2013; Lord et al., 2014; Volker et al., 2010; Volkmar et al., 2014).

For ASD assessments in both clinical and school settings, it is advised that practitioners use: (a) a structured or semi-structured interview with parents, (b) ASD checklists or rating scales completed by caregivers and teachers, and (c) the ADOS-2 or a similar structured ASD observation method (Clark et al., 2014). Most notably, researchers indicate that the most accurate and reliable decisions about ASD school eligibilities and clinical diagnoses come from the use of standardized measures, including information from multiple tools (Aiello et al., 2017; Esler & Ruble, 2015; Lord et al., 2014) Established measures such as the Autism Diagnostic Interview-Revised (ADI-R; Rutter et al., 2003) and the Autism Diagnostic Observation Schedule, Second Edition (ADOS-2; Lord et al., 2012b) are frequently used for diagnostic purposes, as both assess core and associated features present within ASD and are considered "gold standard" instruments for diagnostic purposes. However, such tools require highly trained, proficient professionals for appropriate administration (Lord et al., 2012a) -- and such training is often expensive and time-consuming, which can adversely affect feasibility and adoption. Further, the ADOS-2 and ADI-R were developed to help reliably establish a categorical diagnosis but were not designed to assess clinical features of ASD on a broader continuum or to measure treatment progress, and, as a result, have not been found to be sensitive to small or short-term changes in behavior in individuals with ASD (Bolte & Diehl, 2013; Lord et al., 2014).

Use of Rating Scales in ASD Assessment

Rating scales are another useful method of data collection, as they allow examiners to compare multiple stakeholder perspectives (e.g., parents, caregivers, teachers, relevant community members) across multiple settings, and, when norm-referenced, allow for score comparisons relative to standardization samples (Hughes et al., 2002; Toomey, 2008). Several behavior rating scales and direct observation tools are available to collect information about ASD-related symptoms and behaviors. Tools such as the Gilliam Autism Rating Scale – Third Edition (GARS-3; Gilliam, 2014), Social Responsiveness Scale – Second Edition (SRS-2; Constantino & Gruber, 2012), Childhood Autism Rating Scale – Second Edition (CARS2; Schopler et al., 2010), and Autism Spectrum Rating Scales (ASRS; Goldstein & Naglieri, 2009) are common ASD-specific rating scales cited throughout the literature (Aiello et al., 2017).

Assessment of Restricted and Repetitive Behaviors and Interests (RRBIs)

Restricted and repetitive behaviors have proven to be a difficult construct to measure reliably and validly (Honey et al., 2012; Scahill et al., 2015). Ultimately, none of the previously discussed rating scales captures the broader or full range of the restricted and repetitive behaviors and interests (RRBI) construct—and the field is clearly in need of psychometrically sound instruments that do so. The rating scales that do include subscales and items intended to measure this domain for the ASD population all present with shortcomings in either their nuances of capturing RRBIs, thoroughness dedicated to representing RRBIs, or both. With the absence of RRBI-specific tools, researchers often use other instruments that were not intended to assess nuanced differences in types of RRBIs or, in other cases, not intended to be treatment sensitive—especially for less intensive interventions over shorter periods, such as the ADI-R (Brinkley et al., 2007) and ADOS-2 (Carruthers et al., 2021; Lord et al., 2008).

Currently, five measures specifically assess aspects of RRBIs in the context of ASD. Though they vary considerably from each other in length, scope, and sophistication, several are used regularly in research and some in practice settings. These five instruments are the

Children's Yale-Brown Obsessive-Compulsive Scale for Pervasive Developmental Disorder (CYBOCS-PDD; Scahill et al., 2006), the Stereotypic Behavior subscale of the Aberrant Behavior Checklist (ABC; Aman et al., 1985), the Stereotyped Behavior Checklist (SBC, Rojahn et al., 2000), the Repetitive Behavior Questionnaire (RBQ; Honey et al., 2012), and the Repetitive Behavior Scale-Revised (RBS-R; Bodfish et al., 2000). The listed instruments have all been cited in the literature as significant and appropriate measures of ASD-specific RRBIs (Scahill et al., 2015). Yet, the existing assessments have been examined regarding their validation only based on caregiver reports. With this, there remain questions regarding the reliability, validity, and sensitivity to behavior change over time with the use of teacher or school staff raters. Still, the listed assessments are used in practice for such purposes (i.e., monitoring behavior change) requiring their validation and psychometric support for use with other types of raters (e.g., teachers, school staff, clinical or hospital practitioners, etc.). Of the listed five ASD, RRBI-specific rating scales, the RBS-R stands out for the breadth and nuance of its coverage, and operationalization of ASD-related RRBI subdomains via several explicit subscales (Scahill et al., 2015).

The Repetitive Behavior Scale-Revised (RBS-R)

Before the current, revised edition, the authors published the original Repetitive Behavior Scale (RBS; Bodfish et al., 1999). The original tool was designed to measure the presence of restricted, repetitive behaviors in individuals with ASD (Bodfish et al., 1999). At the time of publication, ASD research in the domain of RRBIs was newly emerging, requiring a tool to assess behaviors in individuals for research and treatment outcome purposes (Bodfish et al., 1999). Scale authors soon revised the RBS (i.e., the RBS-R; Bodfish et al., 2000) by revising and broadening its subscale structure to better differentiate and assess a more comprehensive range of repetitive behavior subdomains. This initial six-subscale structure was based on conceptually derived clusters of items (Bodfish et al., 2000).

To date, the RBS-R has been utilized with increasing frequency, by both researchers and practitioners, to measure changes in RRBIs across time (Boyd et al., 2011; Esbensen et al., 2009; Fulceri et al., 2016; Inada et al., 2015; Scahill et al., 2015; Schertz et al., 2016; Stratis & Lecavalier, 2013; Wolff et al., 2016). Scahill et al. (2015) reported on the psychometric properties of the RBS-R, suggesting that, in the context of parent ratings in ASD samples, the tool shows some reasonable evidence of reliability when used in samples of children, adolescents, and adults (e.g., ages two to over 65 years old). Still, regarding the validity, the RBS-R's factor structure is not a settled issue. Depending on the factor structure ultimately supported, this could affect other prior psychometric evidence (e.g., subtest reliability and concurrent validity estimates depend, in part, on the number of subscales and their item content). Additionally, several psychometric studies have been conducted on translations of the RBS-R in other languages and cultures (Georgiades et al., 2010 [Greek]; Inada et al., 2015 [Japanese]; He et al., 2019 [Chinese]; Kästel et al., 2020 [German]; Martínez-González & Piqueras, 2017 [Spanish]). However, no consensus has been reached regarding the most appropriate and clinically meaningful factor structure for this instrument across both caregiver and teacher raters.

As alluded to above, the psychometrics of the RBS-R have not been evaluated extensively with raters other than parents or caregivers. No study using the original English version of the RBS-R has evaluated the tool's use or factor structure outside of parent/caregiver raters in the ASD context. Thus, what psychometric evidence there is for various types of reliability and aspects of validity such as factor structure, and concurrent relationships with measures of similar or different constructs, is largely restricted to parent and caregiver ratings of individuals with ASD.

The potential of the RBS-R to help address the need for high-quality measures of ASD-specific RRBIs for both intervention research and practice necessitates an investigation of its internal structure validity. Without such validity evidence, the number, and types of subtests on the RBS-R are in question. A noted weakness in the literature on the RBS-R is the lack of statistical evidence to support the author's proposed six-subscale structure in ASD samples rated by parents (Honey et al., 2012; Young & Lim, 2021). As indicated in the section that follows, several other possible factor structures have been identified for the RBS-R using parent ratings of ASD samples. The best-fitting factor model still needs to be identified for this context, but it appears unlikely to be a six-factor model. In contrast, there are no factor analyses of the RBS-R, in the English-speaking context that involve special education staff as raters. This is even though such staff spends considerable time with students who have ASD and need to monitor their progress over time in the school context.

However, it is noteworthy that an EFA of the Spanish version of the RBS-R in Spain involved school staff raters of students with ASD--and it is the only factor analytic study of the RBS-R in the literature to use non-parent raters. These authors (Martínez-González & Piqueras, 2017) retained a six-factor structure from their factor analysis. These results lead to several alternative hypotheses. Was the divergent result due to the different cultural or language context, due to the type of raters involved, or does it reflect mere sampling variation? Though the reason for the finding is not clear, the possibility that using school staff raters instead of parent raters might lead to somewhat different factor solutions is worth examining.

Factor Analyses in the Development of the RBS-R

Guidelines for test development and validation across the literature suggest that the use of factor analytic methods aligns with best practices for evaluating psychological measures (Fabrigar & Wegener, 2012; SEPT, 2014). In general, factor analytic techniques guide researchers to determine the number and basis of factors, or latent variables, which theoretically account for shared variation across different items on a measure. Factors can provide the empirical basis for forming groups of items into scales or subscales. Under the broader domain of factor analysis lies two broad types of approaches: exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). In EFA, the aim is to explore the main constructs or dimensions that may underlie a set of item data, scores from different tests, or other units of analysis (Kline, 1994; Spearman, 1904). A CFA is often used to assess the fit of the hypothesized factor structure (Pett et al., 2003). In that sense, CFA is used as a kind of theory testing, wherein the proposed factor structure is the theory. Accordingly, an EFA is more exploratory in nature while a CFA serves verification purposes regarding the factor structure of an instrument.

At the time of this study, there is a lack of consensus in the literature regarding the most appropriate subscale (i.e., factor) structure of the RBS-R, especially for instrument ratings made by teachers or school staff. All but the one-factor analytic study of the RBS-R used parent raters. This one non-parent rater study examined the most appropriate subscale structure of the Spanish version of the instrument using school staff informants by conducting a CFA following an EFA within the same study. Several studies have examined the structure of the RBS-R using both EFA and CFA methods, though no clear consensus factor structure has been reached across methods. Two EFAs (Bishop et al., 2013; Lam & Aman, 2007) have resulted in a five-factor solution, while one EFA (Russell et al., 2019) has resulted in a four-factor solution for the instrument, all of which differ from the original six-factor structure proposed by the instrument authors. All studies included a sample of individuals with ASD, with ratings provided by participant caregivers.

In the CFA literature, there is more agreement on the factor structure of the RBS-R, when compared to the EFA literature. To date, three studies (Hooker et al., 2019; Mirenda et al., 2010; Sturm et al., 2022), using caregiver ratings have used CFA methods to determine the fit of potential factor solutions for the English version of the instrument. The three studies examined a wide array of factor structures that reflected several conceptualizations of RRBIs in the literature, including an examination of fit for solutions consisting of one- to six factors. However, despite agreement on factor structure, there are some differences in the fit statistics among the studies. One study (Mirenda et al., 2010) yielded support for a five-factor solution with a poorer model fit, compared to the second study's (Hooker et al., 2019) five-factor model fit. Sturm et al. (2022) found a strong statistical fit for a four-, five-, and six-factor model, but landed on a five-factor model based on clinical ease and interpretability.

All three CFA studies (Hooker et al., 2019; Mirenda et al., 2010; Sturm et al., 2022) included only caregiver ratings of preschool-aged and school-aged participants. The inclusion of raters separate from caregivers is important, as context and perspectives have been shown to potentially influence rater behaviors (Tziner et al., 2005). Moreover, whether due to differences in environment/context, education, or experience with ASD and RBBIs raters from a special education background may interpret questions or perceive behaviors differently than caregivers, which could lead to differences in factor solutions across rater types.

Purpose of the Present Study

The purpose of the present study is to examine the factor structure of the RBS-R when samples with ASD are rated by special education staff who know them well. It is possible that special education staff, who are familiar with the characteristics of ASD, may have a more nuanced perception of the RRBIs that they observe in contrast to parents who may not be as familiar with what to look for. Such differences in raters could lead to differences in ratings that, in turn, could result in different factor-analytic findings. As such, a consistent factor structure across different rater types would be a real advantage for comparing scores from different raters. Which possibility is true, is a very important question to answer to properly score and use the RBS-R across different types of raters with whom a child or young person with ASD is likely to regularly interact. The present study addresses three existing gaps in the RBS-R internal structure validity literature. These gaps are (1) a lack of studies of the English language version of the instrument utilizing ratings made by special education staff members in samples of individuals with ASD; (2) a lack of studies exploring alternate factor solutions of the English language version of the RBS-R, which require an identification of the most interpretable and clinically meaningful factor solution; (3) a lack of CFA studies directly comparing all available factor models generated with ASD samples using the English language version of the RBS-R.

This project contributes to the body of internal structure validity evidence for the RBS-R. It consists of two-factor analytic studies—each using an independent sample of students with ASD, ranging in age from three to 21 years old, who was rated by special education staff. The first study involves a thorough EFA of the RBS-R using methods appropriate for ordinal data, assessment of a range of factor solutions based on gold standard factor selection criteria, and use of multiple independent experts to guide factor interpretations and selection of the most interpretable and meaningful factor solution. The second study involves a series of CFAs, followed by factor model comparisons. The CFAs use a robust diagonally weighted least squares estimation procedure to properly account for the ordinal nature of the item data and the likely presence of non-normal distributions. The first CFA examines the absolute and relative fit of the retained factor model from the study one EFA. Additional CFAs examine the fit of all other available factor models for the RBS-R from the current research literature. Once the models are evaluated in terms of fit on their own, they will be compared to determine the model that best fits the inter-item covariance matrix. Nested models will be compared inferentially, while non-nested models will be compared using descriptive information criteria indices.

In the following chapters, this paper explores RRBIs in the context of ASD, assessment strategies, and methods of validating assessment instruments. Specifically, chapter two (Literature Review) illustrates ASD diagnostic criteria from a historical perspective and provides an outline of the conceptualization of RRBIs over time throughout the literature. This review also provides information about ASD assessment instruments and details the use of rating scales to measure core characteristics, specifically in the domain of restricted repetitive behaviors and interests. Further, this chapter explores the purposes of factor analytic procedures and their use in validating rating scales. Specific to the RBS-R, this review provides information about the three different versions of the scale, and evidence from the literature that challenges the RBS-R factor structure suggested by instrument authors (i.e., Bodfish et al., 2000). The third chapter (Method) provides rationales for each study’s research questions, in addition to explicit information regarding the samples and analyses for the two featured studies (i.e., the exploratory and confirmatory factor analyses). The fourth chapter (Results) includes a report and review of the analyses proposed in chapter three. Finally, the fifth chapter (Discussion) provides an extensive description and dialogue of the results from Chapter Four.

CHAPTER TWO: LITERATURE REVIEW

Based on data from the National Health Interview Survey, 1 in 6 children in the United States has at least one of the 10 developmental disabilities tracked by the survey (Zablotsky et al., 2019). Autism spectrum disorder (ASD) is among the 10 DDs tracked by the survey, and according to the Centers for Disease Control and Prevention (CDC), the prevalence of ASD in children by the age of eight years is currently estimated at 1 in 54 (Maenner et al., 2020). The current prevalence of ASD is up from 1 in 59 cases two years ago (Baio et al., 2018) -- and both estimates are up from the CDC's estimates of 1 in 152 from the early 2000s (see Autism and Developmental Disabilities Monitoring Network: CDC, 2007). Overall, at least since the 1990s, the prevalence of ASD has been increasing as our understanding of it and our methods of ascertainment have evolved (Maenner et al., 2020; Newschaffer et al., 2005).

ASD is a neurodevelopmental disorder characterized by impairments in social interaction and social communication, as well as the presence of restricted, repetitive patterns of behavior, interests, or activities (APA, 2013). Though ASD as a diagnostic construct has evolved to some degree since its discovery as a separate syndrome back in the early 1940s (Asperger, 1944; Kanner, 1943) its current two core features have been a part of its core since the beginning. Critically, autism, now ASD, is now conceptualized as a spectrum disorder—with gradations in symptom severity, expression, and associated features that are particular to the individual. Symptoms of ASD are typically present during an individual's early developmental period, though symptoms may not become prominent in those with intact cognitive and language skills until social demands increase at school age. Research indicates that ASD is a lifelong condition that remains stable across the lifespan (Esbensen et al., 2009).

The DSM-5 characterizes or qualifies ASD diagnoses based on the severity of symptoms and the level of functional support required for social communication and restricted, repetitive behaviors (APA, 2013). Severity levels range from Level 1, "requiring support," to Level 3, "requiring substantial support." Individuals assigned to a Level 1 category often have clear impairments in social communication and demonstrate challenges in restricted repetitive behaviors, but require fewer intensive supports (APA, 2013). Often, individuals classified under Level 1 present with difficulty initiating social interactions, organization, and planning, which invariably hinder independence (Weitlauf et al., 2014). Before the most recent iteration of the DSM-5, individuals with ASD characteristics requiring fewer intensive supports were often diagnosed with Asperger's disorder, high-functioning autistic disorder, or pervasive developmental disorder – not otherwise specified (PDD-NOS; Volker et al., 2010). Individuals falling under ASD Level 3 typically present with a limited ability to speak intelligibly and will rarely initiate interactions (Weitlauf et al., 2014). Additionally, individuals with ASD Level 3 support needs face difficulty with verbal and nonverbal expressive communication and often engage in repetitive behaviors (Masi et al., 2017). For the current study, the focus will be on individuals who require substantial support, resulting from severe deficits in social communication and restricted repetitive behaviors.

Autism Spectrum Disorder (ASD) Diagnostic Criteria

With increasing prevalence, practitioners and researchers alike have highlighted the need for more evidence-based, psychometrically sound assessment tools for screening, diagnosis, characterization, and progress monitoring. Research has continuously cited a need for a more uniform assessment process, underscoring validity and efficiency as critical areas of need for both diagnostic and treatment purposes (Delmolino & Harris, 2012; Matson & Kozlowski, 2011). With increasing prevalence rates of ASD, research surrounding the emotional, financial, and familial effects of the disorder has followed suit (Bristol et al., 1996; Horlin et al., 2014; Karst & Van Hecke, 2012; Lavelle et al., 2014) accentuating the importance of correct, early diagnoses for individuals on the autism spectrum.

At present, there are several tools available to assess a range of features and skills in individuals with ASD and/or ID (APA, 2013; Lord et al., 2014; Volkmar et al., 2014). For diagnostic and initial intervention purposes, broad assessments typically include the evaluation of core features, medical aspects, developmental status, cognitive ability, adaptive behavior, speech/language and hearing, and motor skills (Volker et al., 2010). However, the most useful tools designed to inform interventions for individuals with ASD/ID focus on key deficit areas aligned with adaptive behaviors, language, social skills, motor skills, and academic skills that are criterion-referenced and clarify specific needs. Common evaluative tools specific to the domain of ASD diagnosis include observational systems, behavior rating scales, and structured interviews that ask caregivers to reflect on the current and past functioning of an individual (Goldstein, 2018). Despite the format of the assessment tool used, many ASD symptom-related instruments are connected to the DSM-5 (APA, 2013), which is the predominant resource used by clinicians for diagnostic purposes. Given the extended period over which likely participants for the current study were diagnosed with ASD, the project included participant data from those who received ASD diagnoses from either the DSM-IV-TR or the DSM-5. Because of this, relevant diagnostic criteria from both diagnostic manuals are included in the review of the literature.

DSM-IV-TR Diagnostic Criteria

The DSM-IV-TR (APA, 2000) conceptualizes autism theoretically differently from the conceptualization included in the DSM-5. The DSM-IV-TR includes autism-related disorders under the category of Pervasive Developmental Disorders (PDDs). These disorders include childhood disintegrative disorder (CDD) and an umbrella term of five disorders to make up autism spectrum disorder: Rett's disorder, PDD-NOS, Asperger's disorder, and autistic disorder, while childhood disintegrative disorder (CDD) was also included under the broader category of PDDs (APA, 2000). Removal or re-categorization of several of the PDDs occurred with the publication of the DSM-5. A diagnosis of CDD required individuals to have typical development between the ages of four to six years old followed by a sudden regression and appearance of typical autism symptoms (Rutter, 2011). Though now considered a neurobiological disorder due to genetic predisposition, Rett's disorder was initially included in the DSM-IV-TR due to symptom presentation and disintegration of developmental milestones (Volkmar & Reichow, 2013). Additionally, individuals received the diagnostic label of PDD-NOS in the case of the presence of significant symptoms like autism but lacking in areas to have them meet fully the criteria for a PDD (Volker et al., 2010). Similarly, individuals presenting with symptoms of autism and average to above-average skills (e.g., cognitive, language, and adaptive skills) received a diagnosis of Asperger's disorder (Volker et al., 2010). All individuals diagnosed with a PDD (i.e., namely autistic disorder, Asperger's disorder, or PDD-NOS) under the DSM-IV and DSM-IV-TR fit the criteria for ASD with the revision of the DSM-5 (APA, 2013). Largely, research has demonstrated that Asperger's disorder and PDD-NOS are clinically ambiguous and indistinguishable, as clinicians often assigned unreliable diagnostic decisions under the DSM-IV-TR (Lord et al., 2012a; Witwer & Lecavalier, 2008). Upon consideration of research surrounding the PDDs and the participant data for the current study, only core diagnostic information of autistic disorder is featured in this section.

DSM-IV-TR to DSM-5 Changes for ASD

The DSM-5 (APA, 2013) reorganized the DSM-IV-TR (APA, 2000) autism diagnostic criteria into five categorical PDD domains into an overarching category of ASD. With the new conceptualization of the criteria, symptoms related to sensory sensitivity were added, while the remaining criteria were merely reorganized. Among the major shift in disorder conceptualization, major modifications from the DSM-IV-TR (APA, 2000) to the DSM-5 (APA, 2013) included merging three core diagnostic features of ASD into two (i.e., previously social, communication, and restricted, repetitive behaviors to social-communication and restricted repetitive behavior). Furthermore, the DSM-5 requires the use of specifiers to note symptom severity over the previous axial system used in previous versions of the manual. A final general modification includes the allowance of clinicians to assign an ASD diagnosis to individuals who met the criteria historically, rather than with present symptom manifestation. The DSM-5 provides an explanation suggesting that individuals with a DSM-IV diagnosis of autistic disorder, Asperger’s disorder, or PDD-NOS should receive the DSM-5 diagnosis of ASD (APA, 2013). However, individuals presenting with marked deficits in social communication, but with symptoms that do not warrant a diagnosis of ASD may fit the criteria for a diagnosis of social pragmatic communication disorder (APA, 2013; Goldstein, 2018).

DSM-5 Diagnostic Criteria

In the most recent edition of the DSM (i.e., DSM-5), individuals must meet criteria by demonstrating symptoms within the core features: Part A, “persistent deficits in social communication and social interaction across multiple contexts” and Part B, “restricted, repetitive patterns of behavior, interests, or activities" (APA, 2013). Additionally, Part C suggests that symptoms must be present in the early developmental period, though may not be fully discernable until environmental demands surpass limited capacity. Further criteria require that clinically significant impairment must present in important areas of current functioning (e.g., social, or occupational) and disturbances may not be better explained by an intellectual developmental disorder or a global developmental delay (APA, 2013). Evaluators are asked to specify whether ASD presents with or without accompanying intellectual or language impairments, associated with a medical condition (including a neurodevelopmental or behavioral disorder), and with or without catatonia (Goldstein, 2018).

When making a diagnosis, evaluators must also provide levels of severity for the two core criteria domains, including social communication and restrictive, repetitive, behaviors. The severity of ASD is based on increasing intensity levels, ranging from level one, "requiring support," to level three, and "requiring very substantial support" (APA, 2013). Individuals provided with a diagnosis of ASD at level one typically require support to accommodate impairments. Such individuals may present with social difficulties in initiating social interactions and social reciprocity, difficulties with organization and planning, and inflexibility (APA, 2013). Individuals typically diagnosed with ASD in level two require more significant support to help with symptom accommodation and present with social communication difficulties (e.g., limited social interactions with peers, noticeable deficits in verbal and nonverbal communication) and frequently restricted, repetitive behaviors that hinder functioning across environments (APA, 2013). Individuals diagnosed with ASD in level three require the most intensive level of support in place to accommodate presenting impairments. Level three includes severe deficits in social communication skills (e.g., verbal, and nonverbal communication, often little to no intelligible speech) and distressing impairments in restricted, repetitive behaviors (e.g., challenges coping with change, presence of challenging stereotyped or restricted behavior) across all contexts (APA, 2013).

As required for a DSM-5 diagnosis of ASD, individuals must meet the criteria for "persistent deficits in social communication and social interactions across multiple contexts," including the presentation of required behaviors currently or historically. Under the domain of social communication difficulties, individuals must demonstrate "deficits in social-emotional reciprocity," which may range from abnormal social approaches to failure to initiate social interaction (APA, 2013). Individuals must also present with "deficits in nonverbal communicative behaviors used for social interaction," ranging from poorly integrated verbal and nonverbal communication to a total lack of facial expressions (APA, 2013). Additionally, individuals are required to present with "deficits in developing, maintaining, and understanding relationships," varying from behaviors such as difficulties with adapting behavior to fit social contexts to minimal interest in peer interactions (APA, 2013). While all three domains of behavior must meet criteria in some capacity, the examples included are not exhaustive nor necessarily required for diagnostic decision-making.

To meet diagnostic criteria for ASD, individuals must also demonstrate a presence of "restricted, repetitive patterns of behavior, interests, or activities," manifesting through at least two of four listed specific behaviors currently or historically. More specifically, the demonstration of behaviors such as "stereotyped or repetitive motor movements, use of objects, or speech" may contribute to a diagnosis in this domain (APA, 2013). Further, displaying "insistence on sameness, inflexible adherence to routines or ritualized patterns of verbal or nonverbal behavior," falls within this area (APA, 2013). Individuals may also present with abnormally intense or focused “restricted, fixated interests” under the broad category of restricted, repetitive patterns of behavior (APA, 2013). Fourth, exhibiting “hyper- or hyper-reactivity to sensory input or unusual interest in sensory aspects of the environment,” may be present under this broader core area of ASD (APA, 2013). Though not featured as a requirement, the DSM-5 highlights associated features commonly present in individuals with ASD, often considered to fall under the realm of restricted and repetitive behaviors. These behaviors include self-injury (e.g., head banging, wrist biting) and disruptive or challenging behavior (APA, 2013). Additionally, motor impairments and catatonic motor behaviors (e.g., posturing, or grimacing) are frequently associated with the disorder (APA, 2013).

Restricted and Repetitive Behaviors and Interests (RRBIs)

Though RRBIs are strongly associated with ASD, the presence of such behaviors is frequently observed in several additional disorders, as well as in typically developing individuals (Lewis & Kim, 2009). Broadly, repetitive behaviors are characterized by repetition, rigidity, and inflexibility, all of which lack an obvious function (APA, 2013; Lewis & Kim, 2009). Within this domain, behaviors often observed in individuals with ASD include stereotyped motor movements, compulsions, rituals, adherence to routines, insistence on sameness, repetitive use of language, and narrow and circumscribed interests (APA, 2013; Turner, 1999). Restricted and repetitive behaviors are a common part of the phenotypes of other neurodevelopmental disorders, including Rett’s, Fragile X, and Prader-Willi syndromes (Lewis & Kim, 2009). Outside of neurodevelopmental disorders, restricted and repetitive behaviors are often present in Tourette syndrome, obsessive-compulsive disorder (OCD), schizophrenia, and Alzheimer’s disease (Frith & Done, 1990; Nyatsanza et al., 2003; Tracy et al., 1996).

The exhibition of repetitive behaviors intermittently occurs throughout childhood for typically developing children as well. Repetitive motor behaviors (e.g., rocking and flapping) and compulsive, ritualistic behaviors (e.g., insistence on specific food or clothes, rituals) are cited to be more common in early childhood (Leekam et al., 2007; Thelen, 1979). Neurotypical children in the preschool age range tend to exhibit repetitive behaviors that include rigidity around preferences, compulsive ordering or arranging, and ritualization (Evans et al., 1997). Further, typically developing children tend to demonstrate such repetitive behaviors around age two, though the duration and presentation of these behaviors vary on an individual basis (Symons et al., 2005). However, children with ASD tend to exhibit more frequent repetitive sensory-motor behaviors that last for a longer duration than in typically developing children and in children with ID (Arnott et al., 2010; Barber et al., 2012; Szatmari et al., 1989). Some studies have suggested that ASD can be differentiated from other disorders based on the type and severity of demonstrated repetitive behaviors (Bodfish et al., 2000; Mandy et al., 2011), though these findings are not consistent across the literature (Lewis & Kim, 2009).

Patterns of Repetitive Behavior Conceptualization

Historically, RRBI categorization, namely, as it relates to ASD, has varied throughout the literature. Turner (1999) conceptualized these behaviors into two clusters of “higher order” (e.g., compulsions, rituals, insistence on sameness, circumscribed interests, etc.) and “lower order” (repetitive use of objects, stereotyped movements) behaviors, distinct from each other based on their demonstrative complexities. Previous factor analyses of the ADI-R (LeCouteur et al., 2003) have supported this classification of repetitive behaviors, yielding two factors: repetitive sensory-motor behavior and resistance to change/insistence on sameness (Cuccaro et al., 2003; Szatmari et al., 2006). Additional studies have agreed with a two-factor conceptualization, though domain names have varied throughout the literature (Georgiades et al., 2007; Mooney et al., 2009; Richler et al., 2010). Lam et al. (2008) suggested that repetitive behavior should fall within three categories: inflexible language and behavior, repetitive sensory and motor behavior, and circumscribed interests.

RRBIs in ASD

Though the presence of RRBIs is a core feature of ASD, the severity and presentation of RRBIs can differ extensively among individuals. Restricted and repetitive patterns of behavior are often core, early emergent symptoms of ASD that typically present between 18 and 24 months of age (Watt et al., 2008). RRBIs may range from repetitive motor movements to obsessions with parts of objects, to insistence on daily routines for individuals with ASD. Such behaviors may only be observed in specific instances or environmental settings, while some may interfere with adaptive functioning (APA, 2013). Specific to children with ASD, RRBIs tend to manifest early in a child's development and can be detected as early as 17 and as late as 37 months of age (Yirmiya & Charman, 2010). The occurrence of RRBIs varies across individuals with ASD depending on several individual-specific variables, such as age, gender, symptom severity, cognitive functioning, language skills, and adaptive functioning (Hong & Matson, 2021).

In childhood, lower-order behaviors (e.g., stereotyped movements and repetitive use of objects) have been found to remain stable or decrease over time, and higher-order behaviors (e.g., circumscribed interests, insistence on sameness, etc.) have been found to increase over time (Kim & Lord, 2010; Richler et al., 2010). Some research suggests that the presence of ID may affect the trajectory of RRBI topography (Esbensen et al., 2009; Shattuck et al., 2007). Comparably, cognitive functioning has been found to influence the expression of RRBIs.

Positive associations between non-verbal IQ and lower-order repetitive behaviors are present throughout the literature (Bishop et al., 2006; Turner, 1999), whereas support is weaker for associations between nonverbal IQ and higher-order RRBIs (Boyd et al., 2012; Richler et al., 2010).

Regarding ASD symptom severity, several studies have denoted a strong association between the core symptoms of ASD and RRBIs (Dworzynski et al., 2009; Kuenssberg & McKenzie, 2011). One study comparing caregiver ratings on a measure including RRBI components (i.e., the *Aberrant Behavior Checklist*; Aman et al., 1985) and clinician-rated ASD measures (i.e., the ADOS-2; Lord et al., 2012b) provided research support for the association between ASD severity level and topography of RRBIs (Weitlauf et al., 2014). Additional studies including young participants with ASD found that participants with higher levels of ASD severity exhibited more frequent and intense RRBIs than children with lower ASD severity ratings – further suggesting ASD severity has a positive relationship with RRBI frequency and intensity (Hong & Matson, 2021; Matson et al., 2009). Similarly, lower adaptive behavior functioning has been found to be associated with higher RRBIs and sensory behaviors in individuals with ASD (Cuccaro et al., 2003).

Females with ASD have been reported to exhibit RRBIs differently than their male counterparts. Broadly, females have been found to demonstrate fewer and less severe RRBIs, when compared to males with ASD (Bölte et al., 2011; Van Wijngaarden-Cremers, 2014). Further studies have highlighted notable differences, such as less special, narrow, and intense interests (Frazier et al., 2014). Differences in presentations have been thought to be caused by several factors, such as the alteration of brain networks that occur due to gender differences at birth (Langen et al., 2009; Rojas et al., 2009). At this time, much of the literature surrounding

RRBIs and brain anatomy has been studied in males with ASD, leaving room for further explanation of differences at the anatomic level. However, one twin study (van't Westeinde et al., 2020) found within-pair associations between RRBI symptoms and the anatomy of brain networks associated with RRBI symptoms in ASD. Such findings suggest mainly brain structural associations with RRBIs in females and not males, finding increased thickness of the right intraparietal sulcus and reduced volume of the right orbital gyrus in females only (van't Westeinde et al., 2020).

Without intervention, the intensity of RRBIs tends to intensify as children age, often resulting in challenging behaviors (Rispoli et al., 2014). Though not all RRBIs merit intervention, some behaviors tend to interfere with children's learning experiences and academic performance (Koegel & Covert, 1972). Further, highly frequent RRBIs may cause peer exclusion and social isolation (Loftin et al., 2008). Caregivers of children with ASD report RRBIs as a contributor to stress and poor family well-being and functioning (Harrop et al., 2016). The current literature underrepresents interventions specifically targeting RRBIs, compared to interventions targeting social communication. A lack of sensitive screening and assessment measures may play a role in the lack of evidence-based practices to address RRBIs (Raulston & Machalicek, 2018), highlighting the need for a psychometrically-sound RRBI-focused assessment tool.

Assessment of ASD

ASD evaluations require adherence to specific diagnostic criteria (e.g., DSM-5, state special education eligibility criteria, etc.), use of a wide range of assessment tools (e.g., structured observations, rating scales), and consideration of differential diagnoses. The assessment of ASD is complex, as the behaviors often associated with ASD may be due to, or co-occurring with another disorder (Deprey & Ozonoff, 2018). In this vein, assessment of comorbidity and differentiating diagnoses is inherently complex, as children may demonstrate mixed symptoms that stem from a single disorder with an unusual presentation, or one symptom occurring from multiple disorders (Caron & Rutter, 1991; Krueger & Markon, 2006). Specialists involved in both clinical and research-related assessment work must be well informed about symptom presentation and the interaction between mental health issues and other disorders. The key factors that complicate the symptom presentation in cases of ASD include differing developmental levels (Matson & Goldin, 2013), varying cognitive levels (Huerta & Lord, 2012), and the presence of language delays (Lord et al., 2014).

Given the complexities associated with assigning a diagnosis of ASD, evaluators are encouraged to obtain information from both structured and semi-structured interviews, incorporating information from observations across multiple domains. Assessment information should span social behavior, communication patterns, speech, language, intellectual functioning, and restricted or repetitive behaviors (Clark et al., 2014; Ozonoff et al., 2005). Through rating scales or interviews, caregivers provide information about the child's current functioning in adaptive, social, communication, and behavioral areas, in addition to developmental history (Noland & Gabriels, 2004). Examiners should conduct both direct and indirect observations of the individual's social interaction skills (ideally, with same-aged peers) in a naturalistic setting, to better estimate their skills across different environmental settings to complete a comprehensive ASD assessment (Aiello et al., 2017; Seidman & Yirmiya, 2018).

Gathering data from a variety of stakeholders allows for a thorough clinical examination when taken across various methods (i.e., interviews and rating scales; Lord et al., 2014). Caregiver perspectives allow examiners to understand better the individual's developmental history from birth to present, across various contexts and environments (Huerta & Lord, 2012). In addition to caregiver input using rating scales or interviews, teacher input regarding the individual's adaptive classroom behavior, academic functioning, and social and communication should be collected, if obtainable. Such methods, when used in combination with structured and unstructured observations, allow examiners to directly assess the skills and deficits specific to ASD.

Common Observational Instruments in ASD

In diagnostic assessments specifically regarding ASD, examiners must use standardized measures that directly assess the individual's social, communication, play, and behavioral functioning within the context of social situations (Lord et al., 2012a). The inclusion of structured and unstructured interviews and observations to obtain information about the individual's current and historical behavior is essential for correctly identifying the presence of ASD. Recent literature frequently supports the use of the ADOS-2 (Lord et al., 2012b) to assess core areas of impairment associated with ASD (Aiello et al., 2017; Schwartz & Davis, 2014). Additionally, the ADI-R (LeCouteur et al., 2003) is commonly used to obtain information about the child's past and current developmental milestones and functioning across areas of impairment associated with ASD. Together, the ADOS-2 and ADI-R are considered the "gold standard" for assigning a diagnosis of ASD (Falkmer et al., 2013). For both the ADI-R and ADOS-2, clinicians use a protocol of structured or semi-structured interactions or questions to gather more information about the core and associated features of ASD. Protocols are scored according to their respective diagnostic algorithms provided with the assessment tools. At present, the ADOS-2 and ADI-R have the highest sensitivity and specificity to an ASD clinical diagnosis when administered by a clinician who is trained and reliable in the measure (Gotham et al., 2007; Hus & Lord, 2014; Kim & Lord, 2012), highlighting the requisite of advanced training and skill.

Common ASD-Related Rating Scales

While the gold standard combination of the ADOS-2 and ADI-R adheres to best practice guidelines by providing examiners with an opportunity to interview and observe individuals and their parents, it is common practice to utilize rating scales in comprehensive evaluations as supplemental measures (Lord et al., 2014). In the assessment process, rating scales are often used as preemptive screeners, given before, or as part of a comprehensive evaluation (Norris & Lecavalier, 2010a), though they serve a much broader purpose for collecting information across various contexts. Rating scales provide a broader scope that may be more nuanced or contextualized compared to a single observation (Norris & Lecavalier, 2010a; Scahill & Lord, 2004). When compared to many structured and standardized observations and interviews, rating scales are more time- and cost-efficient and require less clinical training (Charman & Gotham, 2013; Scahill & Lord, 2004). Moreover, rating scales allow examiners to collect information and compare different perspectives across multiple settings, in addition to comparing scores to a standardization sample (Hughes et al., 2002; Toomey, 2008).

However, it is important to note that the rater's information is potentially biased in their ratings. The rater is responsible for making a subjective assessment of the individual based on a standardized parameter (Portney & Watkins, 2009) that often does not leave much room for explanation or deviation from the rating scale anchors. Often, the scoring of instruments is influenced by the raters' training and education level, as raters may not have in-depth knowledge and understanding of ASD, diagnostically (Hoyt, 2000). Therefore, raters may not understand the intent or focus of an item, allowing their interpretation to affect outcomes. Additionally, items might include complex language or a higher reading level that increases the difficulty of completing the form (Scahill & Lord, 2004). Best practices in ASD assessments suggest examiners collect information across raters (i.e., parents, teachers, and other caregivers), though ratings performed by multiple raters may result in discrepancies regarding the individual subject's behaviors (Hoyt, 2000). Environmental differences, subjective ratings, or cultural values may all serve as factors that explain different behaviors, or interpretations of behaviors, across contexts (Batista-Foguet et al., 2019; Hoyt, 2000; Martin et al., 2018).

Although many rating scales offer high levels of reliability and validity, potential psychometric limitations include a lack of a representative normative sample for the rated individual and a lack of independent research (e.g., research from authors independent of those who published the scale; Norris & Lecavalier, 2010a). Despite such limitations, practitioners assessing for disorders such as ASD or anxiety favor the use of teacher and parent rating scales over interviews or other forms of assessment (Nathanson & Rispoli, 2022). As with all individual data from assessments, information from rating scales should be used in the context of other assessment tools to allow for multimodal assessment practices (Nathanson & Rispoli, 2022; Salvia et al., 2017).

Common broad-based rating scales used to assess ASD-specific symptom presentations include the GARS-3 (Gilliam, 2014), SRS-2 (Constantino & Gruber, 2012), CARS2 (Schopler et al., 2010), ASRS (Goldstein & Naglieri, 2009), and Social Communication Questionnaire (SCQ; Rutter et al., 2003). Of the listed rating scales, the CARS2 (Schopler et al., 2010) and GARS-3 (Gilliam, 2014) incorporate several best practice guideline elements, such as the inclusion of interviewing raters if needed, adherence to DSM-5 guidelines, and the inclusion of multiple raters for a comprehensive perspective of the individual (Aiello et al., 2017). The SRS-2

(Constantino & Gruber, 2012) and ASRS (Goldstein & Naglieri, 2009) yield strong psychometric properties regarding reliability and validity (Aiello et al., 2017). Though all scales boast unique strengths and weaknesses, multimodal assessment is always recommended, and these rating scales are not intended to be used as the sole indicator of ASD features.

Measurement of RRBIs

Given that the occurrence of RRBIs is a core feature of ASD, it is imperative to assess and monitor their occurrence throughout the lifespan. However, the assessment and monitoring of RRBIs are often complicated by changes in the topography, severity, and frequency of RRBIs across the lifespan (Johnson et al., 2006; Lam & Aman, 2007). As mentioned, several rating scales common to the assessment of ASD are frequently used to measure core diagnostic criteria for the disorder (e.g., the GARS-3, Gilliam, 2014; the SRS-2, Constantino & Gruber, 2012; the CARS2, Schopler et al., 2010; and the ASRS, Goldstein & Naglieri, 2009). Though these rating scales are helpful for screening and assisting (as part of a more comprehensive assessment) with the diagnosis of ASD, they vary in how thoroughly or nuanced they are in representing the core and associated features of ASD. Though several of them represent the social communication domain reasonably well, their coverage of RRBIs is typically more limited or involves combining items that measure RRBIs with items that measure other constructs (e.g., placing RRBI items in a subscale that reflects a broader range of atypical behaviors). Each of the instruments listed above has inadequacies in terms of coverage of RRBIs, and examples for each is provided in what follows.

Though the SRS-2 (Constantino & Gruber, 2012) is described as "an objective measure of symptoms associated with autism" (Constantino & Gruber, 2012, p. 3), and its total score is widely interpreted as measuring general ASD symptoms or traits on a continuum (pp. 27-28), a few considerations should be highlighted. That is, item contents are very heavily weighted toward assessing social communication issues (53 items), while only 12 items are used to assess restricted and repetitive behaviors and interests (see the auto score record forms for the SRS-2). These items do not capture the breadth of RRBIs, and some items are questionable for inclusion among RRBIs (e.g., “Is not well coordinated” or “Talks to people with an unusual tone of voice”). Second, the CARS-2 (Schopler et al., 2010) contains five out of 15 items that, to some extent, assess the realm of RRBIs. However, some of these items include other types of behaviors in the rating scheme and the instrument yields only a total score—with no subscales to reflect different dimensions of ASD or finer distinctions within those dimensions.

In the case of the GARS-3 (Gilliam, 2014), a Restricted/Repetitive Behavior subscale (13 items) combines ritualistic behaviors, stereotypies, sensory issues, and compulsive behaviors. However, some additional items from this domain of ASD are assigned to other subscales intended to assess primarily other constructs (i.e., Emotional Responses and Cognitive Style). Regarding the ASRS (Goldstein & Naglieri, 2009), depending on the age range, contains two or three major scales derived from an EFA that used the combined general population standardization sample and ASD and other clinical cases from the validity studies reported in the test manual. One of these factors, consisting of 23-24 items (again, depending on age), is called "Unusual Behaviors". It is primarily made up of items that reflect stereotypies, inflexibility, and sensory issues. However, it also contains a few items reflecting other types of behavior, such as atypical language use. The authors attempted to break items down conceptually into what they refer to as DSM-IV-TR Treatment Scales, which include three RRBI-related item sets (i.e., Stereotypy, Behavioral Rigidity, and Sensory Sensitivities)—yet these were not supported by their factor analysis.

RRBI-Specific Measures

Researchers often attempt to better understand RRBIs using tools such as the ADOS-2 or ADI-R subscales (Ingram et al., 2008; Klopper et al., 2017). While the ADOS-2 and ADI-R are both psychometrically sound measures, both only provide a unidimensional view of RRBIs, as they both yield one single score. Measures designed to specifically examine RRBIs provide a more heterogenic view, offering a more comprehensive measure of RRBIs overall. It is critical to examine measures of RRBI that are more comprehensive to understand whether they may provide an indication of any ASD subtypes that differ in presentation and outcomes (Zheng et al., 2019).

In a review of measures that specifically assess aspects of RRBIs in the context of ASD, Scahill et al. (2015) highlight several instruments that have been deemed both reliable and sensitive to change over time (Scahill et al., 2015). Though they diverge noticeably from each other in length, scope, and sophistication, most listed instruments are utilized regularly in research settings and practice settings. Measures included in the study were examined for appropriateness regarding their measure of RRBIs. This required meeting the criterion of "good reliability and validity with information on several, but not all, pertinent indices" with "at least one drawback such as only certain subscales are related to repetitive behavior, limited coverage, or available data in a narrow age group" (Scahill et al., 2015). Five measures met this criterion, including the Children's Yale-Brown Obsessive-Compulsive Scale for Pervasive Developmental Disorder (CYBOCS-PDD; Scahill et al., 2006), the Stereotypic Behavior subscale of the Aberrant Behavior Checklist (ABC; Aman et al., 1985), the Stereotyped Behavior Checklist (SBC, Rojahn et al., 2000), the Repetitive Behavior Questionnaire (RBQ; Honey et al., 2012), and the Repetitive Behavior Scale-Revised (RBS-R; Bodfish et al., 2000).

After closer examination, it is evident that all included the validation of these assessments have been examined based on caregiver reports, but there remains a gap in the validation of the instruments using teacher or school staff raters (Scahill et al., 2015). This suggests a failure to provide accurate measurement of important information regarding an individual's skill acquisition, in terms of the scales' usage as progress monitoring tools. Still, the listed assessments are used in practice for such purposes, requiring their validation and psychometric support for use with other types of raters (e.g., teachers, school staff, clinical or hospital practitioners, etc.). Of the listed five ASD, RRBI-specific rating scales, the RBS-R stands out for the breadth and nuance of its coverage, and operationalization of ASD-related RRBI subdomains via several explicit subscales (Scahill et al., 2015).

Rating Variability in Rating Scales

Regarding psychometrics, a principal disadvantage of behavior scales is unintended rating variability. Throughout the literature, rating variability often is attributed to rater variance, setting variance, temporal variance, and instrument variance (Martin et al., 1986 in Campbell et al., 2014). However, rater variability is often cited as a chief concern, highlighting differences across ratings of the same students from respondents in both similar (e.g., home to home, school to school) and different (e.g., home and school) settings (Campbell et al., 2014; Mayes et al., 2009; Sullivan & Riccio, 2007). Rater variance is common among behavior rating scales, boasting low to moderate interrater agreement broadly across scales, though ratings of externalizing symptoms (e.g., repetitive behaviors, aggression) typically reach higher levels of agreement (Achenbach et al., 1987; Gagnon et al., 2007; Winsler & Wallace, 2002).

Rater motive, whether implicit or explicit, can also influence the ratings and reports of behaviors. Stone et al. (2013) found that discrepancies in mother and teacher ratings were more frequent in lower socioeconomic (SES) families, and in some cases, were more pronounced in parents of younger maternal age. Authors (Stone et al., 2013) suggested that parental stress might contribute to reporting of symptom specificity. Specifically, regarding ratings of ASD behaviors, caregivers may be more likely to report their child's behaviors in a certain manner to help their child receive services sooner (Lord & Jones, 2012). On the other hand, caregivers' reports of restricted and repetitive behaviors may be limited if they are afraid to express concerns about their child potentially having ASD (Harstad et al., 2015; Lord & Bishop, 2015; Martin et al., 2018; Schaaf & Lane, 2015).

Setting variance (i.e., variability in behavior dependent on changing contexts) also affects rater responses on rating scales. Respondents familiar with the individual's behavior in one setting may not observe the same behaviors, or topography of behaviors, in another setting. Evidence suggests that interrater agreement between teachers is typically higher than interrater agreements between caregivers and teachers (Achenbach et al., 1987, Brown et al., 2006). The general literature also suggests that parent raters tend to rate their children as exhibiting more interfering behaviors when compared to school staff raters (Culp et al., 2001; Gross et al., 2004; Rescorla et al., 2014; Winsler & Wallace, 2002). Additionally, teachers have been found to rate social withdrawal-related behaviors as more problematic than parents on rating scales of interfering behaviors (Major et al., 2015), suggesting that raters from different settings may prioritize or conceptualize aberrant behaviors differently. Regardless of motive or evidence for reporting differences in behaviors, the general literature underscores the discrepancies between caregiver and teacher-report of behavior.

Rater Influence within Factor Analytic Studies. Given the inconsistencies in rating scale outcomes between caregivers and teachers, there is some evidence to suggest that interrater differences relate to differences in instrument factor structure. In an EFA study that examined the factor structure of the Conners' Parent and Teacher Rating Scales (Conners, 1997) researchers found that teachers and caregivers endorsed items differently. Results of the study yielded significant two- and three-factor models for caregiver reports (rejecting the one-factor model) for caregivers and no appropriate model fit for teacher-reported data (Hardy et al., 2007). Similarly, one study (Bitsika et al., 2016) examining the structure of an ASD-related anxiety and depression checklist (i.e., The Child and Adolescent Symptom Inventory; CASI; Gadow & Sprafkin, 2010) found differences in factor structure regarding reports from caregivers (three-factor) and self-reports (four factors. Though the two solutions were similar, it is noteworthy that the endorsement of items from different raters led to the emergence of an additional factor on the scale (Bitsika et al., 2016). Similarly, two studies (Birnbaum, 2020; Mirwis, 2011) analyzing the factor structure of the Aberrant Behavior Checklist – Community (ABC-C; Aman & Singh, 1994, 2017) using populations of individuals with ASD found different factor structures when using school staff ratings, when compared to the previously established five-factor model of the instrument (Aman & Singh, 1994, 2017).

Monitoring Behavior Change

Though the ADI-R (LeCouteur et al., 2003) and the ADOS-2 (Lord et al., 2012b) are highly considered to be the gold standard tools for diagnostic decision-making in the case of an ASD diagnosis, the instruments are not designed to measure short-term or long-term behavior changes for individuals (Lord et al., 2014). In research and practice, professionals use various assessment tools to monitor behavior change, though few are designed to truly measure behaviors specific to ASD and are not appropriate for measuring a change in this domain (Brinkley et al., 2007). In a systematic review of assessment tools used for this purpose in the field of ASD, McConachie et al. (2015) found several tools with various descriptors of measurement. In this review, most assessments were found to have targeted specific, narrowed categories of behavior related to ASD (e.g., sensory processing, repetitive behavior, challenges in social interaction, etc.), rather than an all-encompassing ASD behavior monitoring tool.

The RBS-R as an ASD Behavior-Monitoring Instrument

The original RBS was developed initially to assess the presence of RRBIs in individuals with ASD (Bodfish et al., 1999). Scale authors cite the need for a scale measuring repetitive behaviors, as research at the time was underdeveloped in the area known to be essential for a diagnosis related to autism or ASD (Bodfish et al., 1999; Bodfish et al., 2000). The RBS-R is a refined version of the original informant-based RBS, both intended to differentiate and assess repetitive behaviors specific to autism spectrum disorders (Lam & Aman, 2007). The RBS-R has frequently been cited as a common practitioner tool for monitoring (RRBI) behavior change across time (Boyd et al., 2011; Chowdhury et al., 2010; Esbensen et al., 2009; Fulceri et al., 2016; Inada et al., 2015; Schertz et al., 2016; Stratis & Lecavalier, 2013; Wolff et al., 2016). The following sections will outline the specific subscales of the most current version of the RBS-R, in the context of the modern conceptualization of restricted, repetitive, and stereotyped behaviors concerning ASD. Of note, the DSM-5 classifies the domain of RRBIs required for ASD into four types (e.g., repetitive, and stereotyped speech, movement of the use of objects; routines, ritual, and resistance to change; circumscribed and restricted interests; hypo- or hyper-reactivity to sensory input, including unusual sensory interests). Conversely, the RBS-R classifies the domain into six types (e.g., stereotyped behavior, self-injurious behavior, compulsive behavior, ritualistic behavior, sameness behavior, and restricted behavior), further described below.

Stereotyped Behavior Subscale. Stereotyped speech, movement, or use of objects is thought to be the first common form of ritualized behavior. The RBS-R does not include items on the Stereotyped Behavior Subscale that reflect repetitive or stereotyped use of language (e.g., repeating favorite sounds, words, sentences, or songs continuously, asking the same question repeatedly), but instead focuses on items on simple motor stereotypies. This is reflected as an area of importance for repetitive behavior measurement, as Turner (1999) noted that stereotyped movements are the most common characteristic of repetitive behaviors. On this scale, Stereotyped Behavior is defined as "apparently purposeless movements or actions that are repeated similarly," including more specific behavior examples at the item level (Bodfish et al., 2000).

Simple motor stereotypies are not thought to be specific to ASD, frequently occurring in typically developing children until approximately age two (Fyfield, 2014). However, movements such as rhythmic body rocking, assorted hand, and arm movements, (e.g., arm-waving, hand flapping), and repetitive pacing, are considered specific to ASD (Gal & Yirmiya, 2021; Schopler, 1995). Relatedly, simple stereotypic movements are more commonly observed in individuals with ASD (88%) than in individuals with other developmental disabilities (61%) overall (Chebli et al., 2016). Past studies (Bishop et al., 2013; Mirenda et al., 2010; Rojahn et al., 2013) have indicated that the Stereotyped Behavior Subscale is significantly correlated with other similar measures, including the Aberrant Behavior Checklist (Aman et al., 1985) Stereotypy subscale, and ADI-R (LeCouteur et al., 2003) Repetitive Behaviors domain total score and CBCL (Achenbach & Rescorla, 2000) Externalizing Problems score. Researchers have also highlighted that these stereotyped movements are often more problematic for observers rather than those displaying the behaviors, due to their degree of inappropriateness in intensity, suggesting their subjectivity and proneness to cultural-oriented judgment (Gal & Yirmiya, 2021).

Self-Injurious Behavior Subscale. Like stereotyped movements, SIBs are not considered unique to individuals with ASD. Namely, SIBs are often associated with two groups: individuals with psychopathologic conditions (e.g., depression, borderline personality disorder, eating disorders, etc.) that primarily involve non-suicidal self-injury and individuals with neurodevelopmental disorders (Sabus et al., 2019). The RBS-R lists "any movement or actions that have the potential to cause redness, bruising, or another injury to the body and that is repeated similarly," as components of SIB (Bodfish et al., 2000). Such behaviors often include but are not limited to, hitting, slapping, biting, scratching, or picking the skin oneself (Bodfish et al., 2000). SIBs are considered non-normative behaviors performed with the intent of physical self-harm but without the intent to die (Crapper & Ernst, 2015). More specific to individuals with neurodevelopmental disorders, the prevalence, and expression of SIBs in children from this population are variable, ranging from approximately 17% to 92% likelihood of individuals with ASD displaying SIBs (Devine, 2014; Fung et al., 2016; Malone et al., 2007; Myers & Johnson, 2007; Richards et al., 2012).

Compulsive Behavior Subscale. Compulsive behavior typically includes any behaviors that are repeated and performed according to a rule or require that tasks be completed in a certain way (Bodfish et al., 2000). Typically, this includes arranging or ordering objects, washing, or cleaning excessively, counting, checking, hoarding, or repeating routine events (Bodfish et al., 2000; Gal & Yirmiya, 2021). Researchers have noted that this behavior is often observable at all ages but emerges most saliently in play for younger individuals with ASD, who may favor playing with parts of toys over the whole toy itself (e.g., spinning car wheels over playing with the car in imaginative fashion; Gal & Yirmiya, 2021). Further, studies (Mirenda et al., 2010; have suggested that compulsive behaviors measured on the RBS-R via the Compulsive Behavior subscale have moderate, significant correlations with the CBCL (Achenbach & Rescorla, 2000) Internalizing Problems score and ADI-R (LeCouteur et al., 2003) Insistence on Sameness domain.

Ritualistic Behavior Subscale. Within the category of the RRBI criterion for an ASD diagnosis within the DSM-5, individuals must present with an adherence to routines, rituals, and/or resistance to change (APA, 2013). As part of the broader category, ritualistic behaviors typically include behaviors that the individual completes during daily living that are performed in a similar manner (Bodfish et al., 2000). Such behaviors are often compulsive in nature, such as an apparent demand for consistency in how individuals act during a specific time (e.g., bedtime, mealtimes, bathroom and dressing, play, social interactions, etc.) frequently (Schopler, 1995). Often, such behaviors cause difficulties in the individual's social environment and may negatively affect their adaptive behavior skills (Gal & Yirmiya, 2021).

Sameness Behavior Subscale. Sameness behavior in the RBS-R refers to an individual's resistance to change or insistence on sameness. At the item level, specific examples include insisting on a particular order of toys or other items, insistence on walking in a particular pattern, sitting in the same spot, resisting a change in activities, and listening to the same music or video (or a piece of it) continually (Bodfish et al., 2000). Individuals with ASD appear to adhere to routines and are not tolerant of changes in routine, requiring a need for structure and predictability throughout the day (Gal & Yirmiya, 2021). In studies that suggest an RBS-R factor model with fewer factors/subscales than the original six-factor model, the Sameness Behavior subscale is often placed into a broader subscale with Ritualistic Behavior (Lam & Aman, 2007; Russell et al., 2019).

Restricted Behavior Subscale. Often analogous to the larger domain of RRBIs, restricted behavior includes a variety of behaviors, actions, and interests. On the RBS-R, restricted behavior is defined as a limited range of focus, interest, or activity within an individual's life (Bodfish et al., 2000). The development of an unusual, narrow, and circumscribed interest in a particular topic is often attributed to ASD (Klin et al., 2007). Such behaviors may include a fascination or preoccupation with a unique subject (e.g., dinosaurs, television shows) part of objects, or objects that move (e.g., fans, clocks; Volkmar et al., 2005). Narrow interests and restricted behaviors are often coupled with impairments in social development, a core diagnostic feature of ASD, creating a level of oddity that prevents the development of peer relations (Cohen & Volkmar, 1997).

Standards for Validity and Test Design

Over time, the emergent need for valid and reliable tools for diagnoses of children with developmental disabilities has been established in both clinical and research domains. As such, educational and psychological assessments serve as important contributions to comprehensive evaluations for ASD and other developmental disabilities (Clark et al., 2014). However, not all tests have been developed properly, leaving room for fault in decision-making about individual diagnoses or treatment planning. As a response to the necessity of psychometrically sound tools, APA, the National Council on Measurement Education, and the American Educational Research Association established the *Standards for Educational and Psychological Testing* (*SEPT*; 2014). The SEPT was developed to provide criteria for the development and evaluation of tests and practices, as well as guidelines for assessing the validity of test score interpretation and intended uses (SEPT, 2014). Broadly, SEPT provides recommendations for cognitive and educational assessments but notes that standards are still useful with other related instruments (SEPT, 2014). Throughout its entirety, the SEPT document outlines standards across a variety of considerations for test use and development. However, examples of standards most relevant to the present studies reflect properties concerning the validity and test design and development.

All SEPT standards are grouped into clusters representing broader areas of test design, use, and development. In the *Validity* section, selected standards in the cluster of *Establishing Intended Uses and Interpretations* accentuate the need for test development to consider the populations in which they will likely be used most. Standards 1.1, 1.2, 1.3, and 1.4 (see Table 1) suggest that using assessments outside of the intended populations or settings is permissible, but should be executed with caution (SEPT, 2014). As suggested, using tests outside of the settings or populations in which they have originally validated calls for expert judgment regarding the interpretability of the tool. If deemed to affect the validity of the instrument, new validity evidence may be warranted to merit its use for others. The cluster regarding *Specific Forms of Validity Evidence* also outlines dependable, recommended methods for examining the validity of an instrument. Specifically, regarding the internal structure of a tool, Standard 1.13 states that the use of multivariate statistical analysis, such as factor analysis, is adequate for demonstrating measurement tool and score validity (SEPT, 2014).

Regarding *Test Design and Development*, the cluster of standards concerning *Item Development and Review* states that test review processes should involve external experts who evaluate the appropriateness of the assessment for use with different populations (i.e., test takers; Standard 4.6, SEPT, 2014). Additionally, SEPT (2014) states that test specifications should be revised with the emergence of new research data or in the face of significant changes regarding the domain represented. As such, the RBS-R (Bodfish et al., 2000) has been in use throughout major changes in the conceptualization of both repetitive behaviors and ASD, broadly.

Table 1. Examples of Standards for Educational and Psychological Testing

Cluster	Standard Number and Description
Establishing Intended Uses and Interpretations	1.1. The test developer should set forth clearly, how test scores are intended to be interpreted and consequently used. The population(s) for which a test is intended should be delimited clearly and the construct or constructs that the test is intended to assess should be described clearly.
Establishing Intended Uses and Interpretations	1.2. A rationale should be presented for each intended interpretation of test scores for a given use, together with a summary of the evidence and theory bearing on the intended interpretation.
Establishing Intended Uses and Interpretations	1.3. If validity for some common or likely interpretation for a given use has not been evaluated, or if such an interpretation is inconsistent with available evidence, that fact should be made clear and potential users should be strongly cautioned about making unsupported interpretations.
Establishing Intended Uses and Interpretations	1.4. If a test score is interpreted for a given use in a way that has not been validated, it is incumbent on the user to justify the new interpretation for that use, providing a rationale and collecting new evidence if necessary.
Specific Forms of Validity Evidence	1.13. If the rationale for a test score interpretation for a given use depends on premises about the relationships among test items or among parts of the test, evidence concerning the internal structure of the test should be provided.
Standards for Test Specifications	4.6. When appropriate to documenting the validity of test score interpretations for intended uses, relevant experts external to the testing program should review the test specifications to evaluate their appropriateness for intended uses of the test scores and fairness for intended test takers. The purpose of the review, the process by which the review is conducted, and the results of the review should be documented. The qualifications, relevant experiences, and demographic characteristics of expert judges should also be documented.
Standards for Test Revision	4.24. Test specifications should be amended or revised when new research data, significant changes in the domain represented, or newly recommended conditions of test use may reduce the validity of test score interpretations. Although a test that remains useful need not be withdrawn or revised simply because of the passage of time, test developers and test publishers are responsible for monitoring changing conditions and for amending, revising, or withdrawing the test as indicated.

Deriving Rating Scales from Factors

A major staple of scale development for psychological instruments includes the use of factor analytic techniques, which are often used for developing and evaluating the psychometric properties of tools (Fabrigar & Wegener, 2012). Within this process, factor analytic techniques aid in guiding the number and basis of factors that are fundamentally present within measurement scales. In rating scales and psychological instruments, factor analyses help researchers better understand more complex instrument property data by taking observed variables (i.e., “factors”) that have similar response patterns, which are associated with a hidden variable (i.e., confounding variable) that is not measured directly. From here, factors are listed according to factor loadings, which indicate how much variation in the data they can explain (Baglin, 2014). Under the broader domain of factor analysis lies two types of processes: exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). Most often, an EFA is more appropriate for researchers looking to develop a preliminary understanding of the factor data and to seek information about the structure of the data, including the number of dimensions present in a set of variables (Baglin, 2014). A CFA is often used for verification purposes, namely, to further examine the structure of data and to confirm the number of dimensions in a set of variables (Brown & Moore, 2012).

Exploratory Factor Analysis (EFA) and Principal Components Analysis (PCA)

Fundamentally, an EFA is a method of distinguishing factors within a dataset, essentially identifying latent constructs within a set of variables (Costello & Osborne, 2005). EFA methods have been used traditionally to explore the possible underlying factor structure of a set of measured variables without the influence of predetermined structures on the measure (Child, 1990). Similarly, a principal components analysis (PCA) has a similar function, as it is typically used as a variable reduction technique, reducing the number of observed variables to a smaller number of components. EFA is often used when there are little or no assumptions about the nature of the relationship in a dataset, as it allows researchers to generate several potential solutions for the factor model (Fabrigar & Wegener, 2012; Osborne & Banjanovic, 2016).

Several variants of the common factor model utilize the principal factor procedures in their estimations (Gorush, 1983), though EFA and PCA differ in the theoretical direction of factors and variables. In PCA, observed variables are grouped to construct the "derived components" (i.e., factors), with each variable contributing a different weight to the derived component (Hatcher & O'Rourke, 2013). The fundamental assumption made in a PCA is that the measured variables are themselves of interest, rather than a hypothetical latent construct (Osborne & Banjanovic, 2016). PCA assumes that all variables are measured without error, meaning that the components retained account for a maximal amount of variance (i.e., total variance) of observed variables, including common variance, unique variance, and random error variance in its calculation of components (Costello & Osborne, 2005). Osborne and Banjanovic (2016) note flaws in this assumption, as analysis may result in overestimated levels of variance in derived factor variables. However, EFA assumes that latent variables give rise to the observed variables, and solutions account for a shared or common variance only. Additionally, EFA allows for error and unique variance to be included in the model, acknowledging the occurrence of imperfect reliability (Osborne & Banjanovic, 2016).

Confirmatory Factor Analysis (CFA)

CFA tests hypotheses that an identified subset of variables legitimately defines a pre-specified factor structure (Gorsuch, 1983), often derived from completing an EFA. CFA of a measuring instrument is most appropriately applied to measures that have been fully developed, including factor structures that have been acceptably validated (Byrne, 2012). In a CFA, researchers indicate which items are meaningful, as well as which items load on which factors, often based on the hypotheses formulated before analysis (Jackson et al., 2009). Researchers can explicitly test hypotheses concerning the factor structure of the data, with the predetermined model specifying the number and composition of factors. These CFA methods differ from an EFA in that the latter, factors are all either correlated or independent, and all items load on every factor (Pedhazur & Schmelkin, 1991). Within a CFA, researchers match observed and theoretical factor structures to determine the goodness of fit of the predetermined factor model (typically from the EFA), allowing for explicit hypothesis testing for factor structure analyses (Gorsuch, 1983).

EFA and CFA as Complements

Both EFA and CFA seek to represent the structure of correlations among measured variables using a set of latent variables. Due to the nature of the procedures used in analyses, the EFA has been considered historically as a theory-generating procedure, as opposed to the theory-testing procedure that CFA generally carries (Stevens, 1996). EFAs are used to investigate data to determine the number of factors that account for the covariation between variables when there are no evident hypotheses about the number of constructs underlying the data. However, researchers have suggested that further investigation must continue past suggesting hypotheses for further research (Muliak, 1972; Nunnally, 1978; Stapleton, 1997). As a complement, CFAs are generally used to advance the process through the evaluation of the strength of the proposed model. However, Fabrigar et al. (1999) note that EFA is a primarily data-driven approach and is more appropriate for use when researchers have a relatively little empirical basis to make strong assumptions about how many common factors exist. Using an EFA in this situation is fitting in that the analysis generates several possible factor models, allowing for consideration of each in a later CFA. In contrast, CFA requires researchers to specify a set number of factors (including pattern specification of zero and nonzero loadings of the variables on the common factors), allowing for the two methods (i.e., EFA and CFA) to be used in conjunction with each other (Fabrigar et al., 1999).

In the literature, some studies have highlighted common faults present within past factor analytic studies (Costello & Osborne, 2005; Norris & Lecavalier, 2010b; Preacher & MacCallum, 2003; Schmitt, 2011). Within the broader realm of factor analyses (i.e., including both exploratory and confirmatory factor analytic procedures), Schmitt (2011) reported that researchers commonly use outdated methods regarding model fit and rotation criteria. Furthermore, Norris and Lecavalier (2010b) reported that such issues were common in the developmental disability literature, specifically regarding the use of EFAs. Some of these issues are present within the factor analysis literature concerning the RBS-R, as the factor structure of the assessment tool was initially conceptually derived (rather than empirically), and consensus has still not been reached regarding the factor structure of the RBS-R.

The Psychometric Foundations of the RBS-R

Throughout the literature, the RBS-R (Bodfish et al., 2000) has been evaluated through several research studies examining the most appropriate factor structure for the instrument. The original RBS (Bodifsh et al., 1999) included four subscales (i.e., Stereotypy, Self-Injury, Compulsion, and Tic Checklists), which were conceptually grouped and decided on by the authors. With the new iteration of the RBS-R (Bodfish et al., 2000), authors included six subscales, but were again, conceptually, and arbitrarily based, as opposed to the use of factor analytic techniques. Lam (2004) reports that Bodfish et al. (2002) performed a principal component analysis (PCA) of the RBS-R, which yielded a six-component structure of the instrument. Since the publication of the RBS-R (Bodfish et al., 2000), several studies have examined the factor structure of the measurement tool, which have produced mixed results that both confirm and differ from the six-factor structure introduced by the authors. Of note, all studies, including those performed by test authors, have been performed with populations of individuals with ASD. The following sections will briefly review the iterations of the RBS-R, including important findings from the related factor analytic studies.

The Repetitive Behavior Scale (RBS)

The first iteration of the Repetitive Behavior Scale (RBS; Bodfish et al., 1999) included four subscales: (1) Stereotypic Behavior, (2) Self-injurious Behavior, (3) Compulsions, and (4) Tics. The RBS was comprised of items influenced by existing scales at the time of development, including the ADI-R (Lord et al., 1994), the Childhood Routines Inventory (CRI; Evans et al., 1997), the Sameness Questionnaire (Prior & MacMillan, 1973), and the Abnormal Focused Affections Checklist (Schultz & Berkson, 1995). No information is available regarding any factor analytic processes or procedures for the establishment of the RBS. The scale was developed using a population of 498 children (ages two to 18 years) with ASD using a statewide mailing survey to the Autism Society of North Carolina. The original study results suggested the scale yielded acceptable psychometric properties, with whole scale Total Score inter-rater reliability estimates of .88 and test-retest reliability of .71 (Bodfish et al., 1995; Lam, 2004; Powell et al., 1996). No information was provided regarding the reliability estimates of the individual subscales.

The RBS – Early Childhood (RBS-EC)

Measure development of the Repetitive Behavior Scale – Early Childhood (RBS-EC; Wolff et al., 2016) was established to fill the need for a repetitive behavior measure for early childhood samples. Items on the RBS-EC were pooled from the RBS-R (Bodfish et al., 2000) and two versions of the Repetitive Behavior Questionnaire (Esbensen et al., 2009 Honey et al., 2012). Scale authors retained and modified items for revision based on clinical experience and consistency with published literature based on behaviors of both typically developing children and children with neurodevelopmental disabilities within the early childhood population ages (Wolff et al., 2016). In addition, new items were developed in response to common repetitive behaviors among infants, toddlers, and preschool-age children with and without neurodevelopmental disorders. Additional items included behaviors surrounding inflexible social routines, visual inspection of objects, restricted movements, and refined categories of motor stereotypy, including 41 items overall (Wolff et al., 2016).

Items on the RBS-EC consist of a five-point Likert-type scale to gauge the frequency of occurrence for each item throughout the past month. Ratings include: 0 – *behavior does not occur;* 1 - *behavior occurs about weekly or less,* 2 – *behavior occurs several times a week*; 3 – *behavior occurs about daily*; 4 – *behavior occurs many times a day*. Item ratings contribute to measures of topography and frequency, which are summed into total or conceptually derived scores. RBS-EC subscales contain measures of repetitive motor, ritual, and routine, restricted behaviors, and self-directed behaviors (Wolff et al., 2016).

Overall, the RBS-EC (Wolff et al., 2016) has established adequate psychometric properties. Regarding reliability, test-retest data yielded Cronbach's alpha estimates all within the acceptable range, ranging from .70 (Self-Directed, six items) to .93 (Repetitive Motor, 9 items). Overall, test-retest data yielded an overall Cronbach's alpha estimate of .90. Regarding validity, the RBS-EC yielded significant correlations with the original RSB (*r* = .43) for the total composite score. For discriminant validity, summary scores correlated weakly (*r* = .17), but statistically significant, with the RBS-EC. Internal consistency of the RBS-EC composite and subscale measures yielded adequate-to-strong correlations (α = .93 for Repetitive Motor [9 items]; α = .75 for Ritual and Routine [10 items]; α = .77 for Restricted Behavior [8 items; α = .70 for Self-Directed [6 items]), with excellent overall internal consistency (α = .90) for the whole measure.

Measures of construct validity for the RBS-EC included a two-stage factor analytic strategy, including both exploratory and confirmatory factor analyses. An EFA was conducted using a maximum likelihood extraction method with oblique (Promax) rotation, using scree plots for factor extraction. A CFA was used to test models derived from the EFA with a cutoff of .35 to determine item retention within models. Model fit statistics included a comparative fit index (CFI), root mean square error of approximation (RMSEA), and root mean square residual (RMSR), with adequate-to-good fit indicated by CFI close to .95, RMSEA < .60, and SRMR <.80 (Wolff et al., 2016).

For the initial EFA sample, the Kaiser-Meyer-Olkin (KMO) index indicated excellent overall sampling adequacy (KMO = .91), and a visual inspection of scree plots suggested either a three- or four-factor solution (Wolff et al., 2016). The four-factor model was like the conceptually derived model, excluding two items loading on the Restricted Behavior factor over the original Ritual and Routine factor. Three models were tested through CFA (e.g., two-, three-, and four-factor models), with test indices indicating acceptable fit for the EFA-derived four- and three-factor models, also suggesting rejection of the two-factor model (Wolff et al., 2016). Wolff et al. (2016) noted that the conceptually derived four-factor model was deemed acceptable in terms of RMSEA and SRMR statistics but results regarding the CFI statistic were just below the recommended range.

The RBS-R

Revision of the original RBS into the RBS-R included several structural changes of the scale, catalyzed by feedback from caregivers and raters following the completion of the RBS in previous studies (Bodfish et al., 2000; Lam, 2004). Feedback from respondents included in the study was used to revise the RBS, including changes regarding ritualized behaviors, insistence on sameness, and restricted interests (Lam, 2004). Bodfish et al. (2002) expanded the original form of the scale by assembling items from existing rating scales that measure aspects of repetitive behavior (e.g., the ADI-R, LeCouteur et al., 2003; the Childhood Routines Inventory, Evans et al., 1997; the Sameness Questionnaire, Prior & MacMillan, 1973; the Abnormal Focused Affections Checklist, Schultz & Berkson, 1995). Authors also reported pulling from past clinical experiences (as reported by Lam et al., 2004). Items on the revised scale, the RBS-R, are grouped conceptually into six subscales, including (1) Stereotyped Behavior; (2) Self-Injurious Behavior; (3) Compulsive Behavior; (4) Ritualistic Behavior; (5) Sameness Behavior; (6) Restricted Behavior. Following the conceptual placement of items into the six factors, Bodfish et al. (2002) examined the factor structure of the RBS-R using a PCA based on a sample of 124 ratings completed by caregivers of children with ASD (reported in Lam, 2004; Lam & Aman, 2007). Based on the PCA, (defined by the scree test, eigenvalues > 1, salient loading > .30, coefficient alpha > .60, and item-total correlations between .2 and .7), findings suggested a six-factor solution (Lam, 2004; Bodfish personal communication with Lam, March 2004). The authors noted that these analyses did not fully support all items to their hypothesized subscale;

however, the six-factor solution was supported, nonetheless. Further, a close examination of the results of the PCA (see Lam, 2004, Appendix A, pp. 98-99) reveals that several items are cross-loaded across multiple factors, indicating that such items require further investigation, analysis, or omission from the scale.

The RBS-R utilizes a four-point Likert Scale ranging from zero to three, conveying the presence and severity of each behavior. Additionally, informants are asked to provide further information about the interference of the behavior using a visual analog scale with anchor labels (see Chapter 3). Like the previous version of the scale, there is no manual nor guidelines for scale use. Omitted recommendations include information regarding the preferred age of the individual, informant characteristics (e.g., the relationship between informant and individual, the capacity to which the informant is familiar with the individual, amount of time the informant has known the individual, etc.), or settings in which the RBS-R should be used. The RBS-R produces raw scores for each of the six subscales and a total score (i.e., Overall Score). There are no standardization samples of the RBS-R, and most psychometric data regarding the instrument has come from independent studies.

Exploratory Factor Analyses of the RBS-R

Although authors have not yet published any research articles that utilize factor analytic techniques of the RBS-R (Bodfish et al., 2000), several studies have examined the factor structure of the scale, both through EFAs (Bishop et al., 2013; Lam & Aman, 2007; Russell et al., 2019) and CFAs (Hooker et al., 2019; Mirenda et al., 2010; Sturm et al., 2022). In addition to the original, English-version of the RBS-R, multiple studies have examined the factor structure of the instrument across various translated versions of the RBS-R in foreign languages

(Georgiades et al., 2010; He et al., 2019; Kästel et al., 2020; Martínez-González & Piqueras, 2017). Details of each study are outlined below. See Table 2 for a summary of the studies.

Bodfish et al. (2002)

It is of note that Bodfish et al. (2002) were cited to conduct a principal components analysis (PCA) of the RBS-R (Bodfish et al. 2000), though specific details are not currently published and are only available cited in other works (i.e., cited in Lam [2004] and Lam & Aman, [2007]). Lam (2004) reports Bodfish et al. (2002) presented the results of the study at a national ASD-related conference and only the findings reported in previously listed works are available currently. Based on Lam's (2004) report, Bodfish et al. (2002) examined the factor structure of the instrument using an exploratory PCA on a sample of 124 ratings completed by caregivers of children with ASD. The findings of these analyses suggested that a six-factor solution is reasonable, as defined by the scree test, eigenvalues > 1.0, salient loading >.30, coefficient alpha >.60, and item-total correlations between .2 and .7 (Lam, 2004). Yet, a close analysis of specific item loadings (included in Lam, 2004: Appendix A) revealed that items included in each factor did not always load most heavily on their assigned subscale. Thus, there remain questions in the literature regarding the most appropriate factor structure of the instrument. While this initial study set the standard for a six-factor model of the newly revised instrument, more detailed information about the participant sample and data analysis methods used would strengthen support for the use of this particular model.

Lam and Aman (2007)

Lam and Aman (2007) were the first to assess the factor structure of the RBS-R, aside from the scale authors, Bodfish et al. (2002; as cited in Lam, 2004; Lam & Aman, 2007). In their study, Lam and Aman (2007) highlighted the importance of the scale, as it emerged as the first scale assessing specific repetitive behaviors for individuals with ASD but noted the small sample size used in the Bodfish et al. (2000) study as a limitation for determining factor structure. The authors also highlight the significance of placing items into subscales, which was not completed in the previous study. Study authors note that an EFA was used over a CFA due to the lack of empirically driven EFA techniques used to create the RBS-R (Lam & Aman, 2007) previously.

For this analysis, Lam and Aman (2007) sampled parent and caregiver (of children with ASD) responses on the RBS-R, including 307 participants. Participant data included individuals with ages ranging from three to 48 years or (*M* = 15.34, *SD* = 9.60) and 52 females (16.9%). Most of the sample identified as White (69.1%), also including those identifying as Black (23.1%), Asian (2.6%), Hispanic (1.3%), and “Other” (2.3%). All participants had been sampled through the South Carolina Autism Society mailing list, with most participants reporting a diagnosis of autism (81.4%) and severity of ASD reported as “equally divided” across the three severity levels (Lam & Aman, 2007). Within this sample, 16.6% attended general education placements for school, 57.7% attended general education school but received special education services, and 14.7% attended a special school (Lam & Aman, 2007).

Prior to conducting factor analytic processes, the authors first concluded that there were no items to be dropped, based on the study sample’s item endorsement (i.e., authors examined item-level responses to calculate any potential responses that were too high or too low, by collapsing severity ratings, Lam & Aman, 2007). An EFA was then conducted using the inter-rater correlation matrix using items from the RBS-R. This study (Lam & Aman, 2007) utilized the Ordinary Least Squares (OLS) discrepancy function in Comprehensive Exploratory Factor Analysis (CEFA) with oblique quartimax rotation, reporting the use of the OLS due to the potential of the scale constructs compromising repetitive behavior in this specific population.

The number of factors to retain was guided by the scree plot method (Cattell, 1966), eigenvalues greater than 1.0, goodness-of-fit as estimated RMSEA (Browne & Cudeck, 1992), and interpretability. All items were considered to load on a factor if they reached the threshold of loading .35 or higher, in addition to the loading being at least .10 higher than the loading on any other factor. With this criterion, five items were omitted from the scale.

Lam and Aman (2007) examined solutions between two and six factors. Based on set criteria for retention, results yielded two solutions that fit all criteria, which included four- and five-factor solutions. The five-factor solution was chosen over the four-factor solution due to the inclusion of the fifth factor, Restricted Interests, as it was explained to have clinical significance for individuals with ASD (Lam & Aman, 2007). Factors retained included (I) Ritualistic/ Sameness Behaviors, (II) Self-Injurious Behaviors, (III) Stereotypic Behaviors, (IV) Compulsive Behaviors, and (V) Restricted Behaviors. The EFA for the five-factor solution yielded an RMSEA of .061, placing the model fit between reasonable ($\leq$.08) and poor ($\geq$.10; Browne & Cudeck, 1992 in Lam & Aman, 2007). The mean factor loadings for factors one through five were .55, .60, .51, .53, and .53, respectively, with the five factors accounting for approximately 47.3% of the variance (Lam & Aman, 2007).

The most notable difference between Lam and Aman’s (2007) proposed factor solution and the original six-factor structure of the RBS-R (Bodfish et al., 2000) is that the five-factor solution collapsed the original Ritualistic Behavior and Sameness Behavior subscales into a broader Ritualistic/Sameness Behavior subscale. Lam and Aman (2007) note that even when six factors were extracted, the two subscales continued to load on one singular factor. Conceptually, this relationship fits the concept that rituals are strongly related to the need for sameness and consistency. The authors also note that on the remaining four subscales of the proposed five-factor (subscales) version of the RBS-R, 22 of the remaining 26 items (85%) were retained as loaded on the original version of the RBS-R.

Bishop et al. (2013)

Bishop et al. (2013) performed an EFA with an ASD population to examine the presence of more specific behaviors (e.g., Insistence on Sameness (IS) behaviors and Repetitive Sensory Motor (RSM) behaviors) within the broader domain of restricted and repetitive behaviors. Though the factor analysis was not the primary focus of this study (Bishop et al., 2013), the authors provide detailed information regarding the factor structure of the RBS-R, including a large sample size of school-aged children with confirmed ASD clinical diagnoses. Within the analyses, Bishop et al. (2013) compare findings among the presence of IS and RSM behaviors between set intelligence quotient (IQ) and gender-mediated groups of participants, providing phenotypic data to examine relationships and presentation differences.

Participants in this study were recruited from the Simons Simplex Collection, a genetic study that includes families with one child with ASD without any first-, second-, or third-degree relatives with the disorder (Fischbach & Lord, 2010 in Bishop et al., 2013). Bishop et al. (2013) sampled 1,825 individuals with a clinical diagnosis of ASD ranging in age from four to 18 years old (M = 8.9, SD = 3.5). All diagnoses were confirmed through comprehensive evaluations including the ADOS (Lord et al., 1999), ADI-R, Vineland Adaptive Behavior Scales, Second Edition (VABS-II; Sparrow et al., 2005), and cognitive assessments. The sample included 79% White participants, 11% Hispanic participants, and 86% male participants with a mean full-scale IQ (FSIQ) of 83.0 (SD = 26.9). Additionally, the mean verbal IQ was 79.8 (SD = 30.2), the mean nonverbal IQ was reported as 86.5 (SD = 25.1), and the mean VABS-II composite was reported as 74.1 (SD – 11.7). Children with significant hearing, vision, or motor problems or significant early medical histories were excluded from the study. Participant caregivers completed all RBS-R ratings.

The EFA was conducted using promax rotation via Mplus (Muthen & Muthen, 1998-2017). A cutoff of .30 was selected for including an item on a factor, requiring the loading to be at least .10 higher than the loading on other identified factors. Results of the study list RMSEA and chi-square values as indicators of solution fit to the data, citing Hu and Bentler (1999) as the standard for appropriate fit. Authors (Bishop et al., 2013) do not list the full range of factor solutions yielded resulting from this study, but site past EFA studies examining the RBS-R factor solution, including ranges of two to six factors (e.g., Georgiades et al., 2010; Lam & Aman, 2007).

Bishop et al. (2013) list endorsement of a five-factor solution based on fit to the data using the listed criteria above (chi-square (698) = 2202.42; RMSEA = 0.05). Factors retained in the Bishop et al. (2013) study included (I) Sensory-Motor Behavior, (II) Restricted Interests, (III) Self-Injurious Behavior, (IV) Compulsive Behavior, and (V) Ritualistic/Sameness Behavior. The proposed factor solution was similar to the solution proposed by Lam and Aman (2007), though some discrepancies were present regarding item placement. When compared to the original RBS-R factors/subscales, Factor I was similar to the original Stereotypy subscale, Factor II included two out of four items from the original Restricted Behavior subscale, and Factor III was identical to the original Self-Injury subscale (Bishop et al., 2013). Furthermore, Factor IV was "almost identical" to the original Compulsive subscale with one item omission and two item additions from the original Sameness and Ritualistic subscales (Bishop et al., 2013). Factor V included all original items from the Sameness subscale and two additions from the Ritualistic subscale (Bishop et al., 2013).

Overall, Bishop et al. (2013) provided a few important contributions to the RBS-R literature with this EFA study. First, the addition of unique factor names regarding sensory-motor and insistence on sameness behaviors lend for future comparisons between gold-standard instruments, such as the ADOS-2 and ADI-R. Second, the authors highlighted that their five-factor solution was similar to other factor solutions found in a previous EFA for individuals with ASD (e.g., Lam & Aman, 2007). However, scores on the proposed factor solution by Bishop et al. (2013) yielded high inter-factor correlations, which authors suggest may provide some evidence for a three-factor model of the RBS-R as proposed by Mirenda et al. (2010) in a CFA study.

Russell et al. (2019)

Like Bishop et al. (2013), Russell et al. (2019) used the RBS-R to examine further the relationship between specific subtypes of RRBIs associated with ASD. Namely, Russell et al. (2019) sought to examine the relationship between repetitive behavior subtypes and anxiety, influenced by ASD literature at the time (Joseph et al., 2013; Stratis & Lecavalier, 2013) which suggested an interaction between the two sets of behavior presentations. Authors (Russell et al., 2019) site utilization of factor analytic methods in their study due to the lack of agreement on a factor structure of the RBS-R. Additionally, the EFA of the RBS-R was theorized to help with the identification of meaningful subtypes of repetitive behavior and to strengthen the exploration of a relationship with anxious behaviors in children with ASD.

For these analyses, Russell et al. (2019) utilized data collection methods similar to those used in the Bishop et al. (2013) EFA study. Russell et al. (2019) sampled 2,093 individuals with ASD, recruited from the Simons Simplex Collection (SCC). Participants were majority White (79.1%) and male (86.6%), ranging in age from 68 to 216 months (*M* = 123.50, *SD* = 37.52). The average FSIQ for the sample was 81.55 (*SD* = 28.48). All participants were recruited through local service providers, parent advocacy groups, and advertisements (collected in Fischbach & Lord, 2010), and received confirmed clinical ASD diagnoses using the DSM-IV-TR criterion via the ADOS (Lord et al., 1999) and ADI-R. Parents completed ratings of participants on the RBS-R, in addition to the Child Behavior Checklist (CBCL; Achenbach & Rescorla, 2000).

An EFA was conducted with principal axis factoring and promax rotation. To determine the number of factors to retain, the authors used the eigenvalues-greater-than-one criterion, examination of the scree plot, and subjective clinical interpretability. Prior to analysis, the authors set the criterion that all items loading at least .35 or higher and at least, .10 higher than the loading on any other factor would be retained. Additionally, Pearson correlations were used to examine the relationship between repetitive behavior subtypes and anxiety symptoms, as rated by caregivers on the CBCL (Achenbach & Rescorla, 2000).

Russell et al. (2019) presented several solutions, including one-, two-, three-, four-, five-, and six-factor solutions, similar to the literature regarding the factor structure of the RBS-R (i.e., Bishop et al., 2013; Lam & Aman, 2007). The authors highlight similarities between their factor solution data and past EFA literature on the RBS-R, providing evidence for a five-factor solution. However, data from this study (Russell et al., 2019) only supported two items loading on the fifth factor, violating the recommendation of a three-item minimum per factor (Costello & Osborne, 2005, as cited in Russell et al., 2019). Therefore, a four-factor solution was selected with the omission of eight items from the scale overall due to low factor loading. The final four factors included (I) Rituals/Sameness, (II) Compulsive Behaviors, (III) Stereotypic Behaviors, and (IV) Self-Injurious Behaviors. Each factor yielded adequate internal consistency, with alpha coefficients $\geq$ .70. Regarding variance, each factor (I – IV) was listed to explain 23.56%, 5.00%, 3.08%, and 2.19% of the variance, respectively.

The findings from the EFA were generally consistent with previous factor analyses, with items from the original (Bodfish et al., 2000) Ritualistic Behaviors and Sameness Behaviors subscales loading on a single factor, like in Lam and Aman (2007) and Bishop et al. (2013) studies. In this analysis, there was no clear Restricted/Circumscribed Interests factor, which is often considered an associated feature of ASD. Authors (Russell et al., 2019) hypothesize that this may be due to the lack of items representing restricted and circumscribed interests in the RBS-R itself. Additionally, the authors hypothesized that analysis results would suggest a relationship between the Sameness Behaviors and Compulsive Behaviors subscales on the RBS-R and anxiety-related symptoms. However, no independent relationships were found after controlling for other repetitive behavior factors. While this study provides specific information regarding the relationship between anxious behaviors and repetitive behaviors in individuals with ASD, it also provides further complications in setting a unified factor structure of the RBS-R.

Table 2. Summary of Exploratory Factor Analyses of the Repetitive Behavior Scale-Revised (RBS-R)

Study	Source	*N*	Sample Demographics	Sample Diagnoses	Rater	Factor Analysis/ Factor Retention Method	Factor Solution Chosen	Factor Solution Names
Bodfish & Lewis (2002), cited in Lam (2004)	Autism Society of North Carolina	124	N/A	Individuals with ASD	Parents	Principal Components Analysis (PCA)	6	I: Stereotyped Behavior II: Self-Injurious Behavior III: Compulsive Behavior IV: Ritualistic Behavior V: Sameness Behavior VI: Restricted Behavior
Lam & Aman (2007)	Autism Society of America	307	82.4% male; 69.1 % White; 1.3 % Hispanic; Mean age: 15.34 (9.60)	Autistic disorder 81.4%; Asperger's 6.5%; PDD-NOS 6.2%; None reported 4.6%	Parent/ primary care-giver	EFA using inter-item correlation matrix with OLS discrepancy function in CEFA with oblique quartimax rotation Used scree plot method, Eigenvalues above 1.0, RMSEA Goodness of fit	5	I: Rituals/ Sameness II: Self-injurious Behavior III: Stereotypic Behavior IV: Compulsive Behavior V: Restricted Interests

Table 2 (cont'd)

Bishop et al. (2013)	University -based sites	1,825	86% male; 79% White; 11% Hispanic; Mean FSIQ 83.0 (26.9); Mean age: 8.9 (3.5)	All participants received a clinical diagnosis of ASD or autism	Parents	EFA using promax rotation Used chi-square and RMSEA	5	I: Sensory-Motor Behavior II: Restricted Interests III: Self-Injurious Behavior IV: Compulsive Behavior V: Ritualistic/ Sameness Behavior
Russell et al. (2019)	Research database	2,093	86.6% male; 79.1% White; Mean FSIQ: 81.55 (28.48); Mean age: 123.50 (37.52) months	Children with ASD	Parents	EFA to determine which subscales to include in the regression model (with PAF and Promax rotation) Used Eigenvalues above 1, scree plot, and interpretability for factor solution	4	1: Rituals/Sameness II: Compulsive Behaviors III: Stereotypic Behaviors IV: SIB

Confirmatory Factor Analyses of the RBS-R

See Table 3 (below) for a summary of the studies.

Mirenda et al. (2010)

Mirenda et al. (2010) performed a CFA with an ASD population to assess the factor structure of the RBS-R in response to the lack of consistency in the current literature at the time of publication. Within the study, Mirenda et al. (2010) proposed adversarial considerations of methods used by Lam and Aman (2007), regarding participant sampling methods and diagnostic confirmation of individuals included in the study. The authors also emphasize the need for clarification of the RBS-R with increased use of the instrument for both research and clinical purposes. Beyond the CFA performed, another purpose of this study was to examine the external validity of the RBS-R through the examination of correlations between the RBS-R factors and various subscales of additional scales, including the ADI-R, ADOS, Vineland-II, and Merrill-Palmer-Revised Scales of Development (Roid & Sampers, 2004).

For the study, Mirenda et al. (2010) sampled 287 preschool-aged children with ASD (84.3% male, 71.8% White) ranging in ages from two to four (M = 3.33, SD, = .78). Participants were recruited through a Canadian multi-site longitudinal study for children with ASD and were required to have a clinical diagnosis of ASD (using DSM-IV criteria) confirmed by the ADOS and ADI-R. Exclusion criteria consisted of any physical (i.e., vision or hearing), neuro-motor disorders (e.g., cerebral palsy), or genetic disorders. Participants' caregivers completed all ratings on the RBS-R.

CFA procedures were used to test six structural models of repetitive behavior (i.e., beyond the RBS-R, though still including models proposed in EFAs of the RBS-R) present in the literature at the time. It is of note that the tested five-factor model was similar to that of Lam and

Aman (2007)'s model, though Mirenda et al. (2010)'s model included all items, while Lam and Aman (2007) omitted several items due to low factor loadings. Tested models ranged from one to six factors, evaluated using five statistical indices: model chi-square/degrees of freedom, CFI, the Tucker-Lewis Index, RMSEA, and the SRMR to test the goodness of fit of each model to the data. Values less than 3.0 for the chi-square, greater than .90 for the CFI and TLI, below .08 for the SRMR and values of .06 or less were used for indicating model fit for each associated method.

Results of the CFA suggested that the three-, four-, five- and six-factor models were all "reasonably good" fits for the data, as determined through criteria previously discussed (Mirenda et al., 2010). Ultimately, the authors based their model preference according to fit statistics and parsimony standards, ultimately highlighting the three- and five-factor models. The three-factor model includes factor/subscales (I) Compulsive Ritualistic Behaviors, (II) Self-Injurious Behaviors, and (III) Restricted Stereotyped Behaviors; while the five-factor model includes subscales (I) Stereotyped Behaviors, (II) Self-Injurious Behaviors, (III) Compulsive Behaviors, (IV) Ritualistic Sameness Behaviors, and (V) Restricted Behaviors. Between the two proposed models, key differences include the unification of the Compulsive Behaviors and Ritualistic Behaviors subscales in the five-factor model and the amalgamation of Stereotypic Behaviors and Restricted Behaviors in the three-factor model. Overall, Cronbach's alpha estimates were at least .72 or higher for all factors in all six models, an indication of adequate or better internal consistency.

The fit indices reported for each model within this study provide evidence for claims regarding the dimensionality of repetitive behavior in general. Of all tested models, the unidimensional (i.e., one factor-model) yielded the least acceptable fit (RMSEA value of .082), suggesting a more dimensional structure of repetitive behaviors for children with ASD in general (Mirenda et al., 2010). The fit index results indicate that the five- and six-factor models were almost identical, with the five-factor model preferred for parsimony. Authors note that the three- and four-factor models, while not chosen as the best fit, still present acceptable fit statistics, and may still be useful in some circumstances requiring broad descriptors of repetitive behavior (Mirenda et al., 2010). However, it is noteworthy all models highlighted by the authors (e.g., three-, five- and six-factor models) all yielded fit index estimates that were either near or below their respective cutoff levels. Specifically, CFI and TLI estimates were not consistent with best practice guidelines, while RMSEA and SRMR estimates just barely met the recommended cutoffs.

Hooker et al. (2019)

Hooker et al. (2019) performed a CFA of the RBS-R in an attempt to clear some inconsistencies regarding the factor structure of the instrument. Authors call for further analyses of the RBS-R to characterize properly the psychometric properties for the measure for uniformity purposes. Hooker et al. (2019) list issues such as large age range in past studies (e.g., Bishop et al., 2013; Georgiades et al., 2010; Lam & Aman, 2007) as potential masks of quantitative and qualitative differences regarding repetitive behavior across groups. The primary function of the study was to strengthen the understanding of the factor structure of the RBS-R. Beyond this, a secondary goal was to provide information regarding the performance of the RBS-R concerning diagnostic domains for use with clinical samples using the ADOS (Hooker et al., 2019).

Participants in the study included 350 children with ASD, ranging in age from two to nine (*M* = 4.56, *SD* = 2.13) years old. The sample consisted of 82.9% male, 70.6% White, and 16.6% Hispanic participants. All participant diagnoses were confirmed with the ADOS, in addition to the use of the Vineland-II and a cognitive measure (i.e., Mullen Scales of Early Learning or Stanford-Binet Intelligence Scale). Participant recruitment was drawn from two larger studies – one from a large cluster-randomized trial involving a comprehensive engagement of children with ASD in the classroom, and the other an ongoing speech and play project for children with ASD.

For the CFA, Hooker et al. (2019) utilized a robust weighted least square estimation (WLSMV) due to the use of ordinal variables and multivariate normality of the data used. Regarding model fit indices, researchers used the RMSEA, CFI, TLI, and WRMR methods. Hooker et al. (2019) used the maximum likelihood (ML) cutoffs suggested by Hu and Bentler (1999) for the WLSMV method, and values less than .06 for the RMSEA were used as the cutoff. Additionally, CFI and TLI values greater than .95 were considered an adequate fit. In the final model tested, the authors examined item factor loadings, factor correlations, and the residual correlation matrix for model fit (Hooker et al., 2019). Additionally, two indices of fit were extracted, including the Akaike Information Criteria (AIC) and the sample-size-adjusted Bayesian Information criterion BIC. The model composition was based on previous EFAs of the RBS-R (Bishop et al., 2013; Lam & Aman, 2007; Mirenda et al., 2010), the original RBS-R (Bodfish et al., 2000), the DSM-5 (APA, 2013) RRBI criteria, and repetitive behavior subtypes in the literature (Leekam et al., 2011).

Hooker et al. (2019) examined a one- through six-factor solution for the CFA. It is of note that the five-factor model tested in this study was identical to the model from Mirenda et al. (2010). Across the models, Model I (unidimensional factor model) provided the poorest fit to the data, while Models II-VI yielded reasonable fit values. Models V and VI demonstrated the best-fit values, with the most support provided for Model VI, the six-factor model. Model V (five-factor model) was ultimately selected as the most parsimonious model. Besides simplicity, the authors list the selection of the five-factor model due to a good fit and low AIC and BIC values. Further, this model collapsed the "sameness" and "ritualistic" factors from Model VI, which yielded a factor correlation of .88 in the six-factor model – suggesting a significant overlap between the two factors (Hooker et al., 2019). Authors note that there was some evidence to support the four- and six-factor models as well, speculating that differences in data may be due in part to the utilization of the WLSMV method over other methods.

Sturm et al. (2022)

In a more recent study examining the factor structure of the RBS-R, Sturm et al. (2022) performed a CFA of the instrument. Sturm et al. (2022) discuss the importance of prioritizing the calculation of a total score on the RBS-R, stating that existing studies focused more on the scale structure and item retention, rather than a total score calculation. Further, authors (Sturm et al., 2022) cite the need for more rigorous psychometric studies concerning the RBS-R. The authors cite the utilization of more item response theory (IRT)-focused methods to evaluate rigorously the item level, subscale, and whole-scale properties. Methods were also aimed to focus on reliability, validity, and item fit within the scale.

This study sampled participants from three large existing ASD data repositories; Simon Foundation Powering Autism Research for Knowledge (SPARK; Feliciano et al., 2018), SSC (Fischbach & Lord, 2010), and National Database for Autism Research (NDAR). Participants included 15,318 individuals with a mean age of 9.24 (*SD* = 4.0) and a mean non-verbal IQ of 78.4. Further, 17.6% of the sample had an additional diagnosis of intellectual disability. The sample majority was male (80.9%), White/Caucasian (73.2%), and not Hispanic (78.5%). All participants had a diagnosis of ASD, though neither procedures nor assessment batteries used to obtain the diagnoses were not available for the entire sample. For participants pulled from the SPARK database, authors report that 60.7% of the sample used longer sentences, 16.7% were able to combine only three words, 11.6% used single words only, and 11.0% were non-verbal.

Prior to conducting the CFA, Sturm et al. (2022) also conducted an EFA, citing inconsistent results from past studies. The authors conducted the EFA with target rotation using flexMIRT version 3.5 (Cai, 2012), followed by a CFA. The authors divided the sample by language level and ran analyses separately for each sample group. Following EFA with target rotation, iterative CFA models were run to test for weak loading times on each factor individually. Authors used EFA with target rotation and iterative CFA models to guide factor assignment for weak loading items (Sturm et al., 2022). Listed factor assignment criteria included IRT model fit indices, factor loadings, and factor interpretability. Sturm et al. (2022) only list EFA and confirmatory model results for the model that they selected (i.e., the five-factor model). The final model did not permit cross-loadings for model parsimony and the authors state that this was acceptable due to the nominal impact of the weak items on IRT scores. Confirmatory models were fit using Mplus Version 8 (Muthén & Muthén, 2017) using the pooled sample. Methods included the use of CFA for categorical indicators using a robust WLS estimator. Factor variances were set to 1, and authors considered fit indicators by examining model fit indices (i.e., RMSEA, CFI, TLI, SRMR), factor interpretability, factor loadings, and potential subscale or total scores to be created from the model (Sturm et al., 2022).

As stated, authors do not share the results for all models tested, but Sturm et al. (2022) state that of the five CFA models tested, only the unidimensional model (i.e., one-factor model) yielded an inadequate fit by acceptable standards. Results from the study produced similar fit statistics for the four-, five-, and six-factor models, but authors ultimately chose the five-factor model due to its superior model fit (Sturm et al., 2022). A second-order factor model was also explored, but the authors eliminated this model due to its high level of shared variance across the subdomains. The authors also highlight the research evidence from past studies (Bishop et al., 2013; Hooker et al., 2019) and the inclusion of a separate factor to represent restricted interests that is missing in the four-factor model. The proposed model includes (I) Stereotyped Behavior, and (II) Self-Injurious Behavior, while specific labels for Factors III-V are not provided. The fifth factor is described as incorporating "restricted/circumscribed/special interests (Sturm et al., 2022). The internal consistency reliability of the five-factor model was evaluated and reported as Cronbach alphas, indicating moderate (α_{F1} =. 79, α_{F2} =. 76, α_{F3} =. 74, α_{F4} =. 88) to low (α_{F5} =. 66) internal consistency across the dimensions.

Like past studies (e.g., Bishop et al., 2013; Lam & Aman, 2007; Mirenda et al., 2010), this study provided further evidence of the RBS-R as a multidimensional instrument, leaving minimal support for the use of a total score in clinical practice. Compared to past studies examining the factor structure of the RBS-R using caregiver ratings, this study utilized the most advanced and sophisticated analyses to date. However, even with the inclusion of advanced IRT-focused analyses, this study also found very similar model fits for the four-, five-and six-factor models, comparable to past studies. This suggests that there is no one, clear, strong data-emergent model.

Table 3. Summary of Confirmatory Factor Analyses of the Repetitive Behavior Scale-Revised (RBS-R)

Study	Source	*N*	Sample Demographics	Sample Diagnoses	Rater	Factor Analysis/ Factor Retention Method	Factor Solutions Examined (Supported)
Mirenda et al. (2010)	Research Institute	287	84.3% male; 71.8% White; Mean Developmental Index age 23.1 (11.4) months, Mean age: 40.7 months (9.3)	All participants had a clinical diagnosis of ASD	Parents	Confirmatory Factor Analysis (CFA) based on constructed five-factor model Chi-square/degrees of freedom, comparative fit index (CFI), Tucker-Lewis's index (TLI), root mean square error of approximation (RMSEA), and Standardized Root Mean Square Residual (SRMR); Akaike Information Criteria (AIC)	1-6 (5)
Hooker et al. (2019)	Healthcare & childcare agencies; Elementary school	350	82.9% male; 70.6% White; 16.6% Hispanic; Mean IQ Early Learning 74.62 (22.64), Abbreviated IQ 72.98 (19.43); Mean age: 4.56 years (2.13)	Children with clinical diagnosis of ASD	Parents	CFA (robust weighted least squares estimation; WLSMV) Model fit: RMSEA, CFI, TLI, and the weighted root mean square residual (WRMR)	1-6* (5) *Utilized Mirenda et al. (2010) five-factor model

Table 3 (cont'd)

Sturm et al. (2022)	Several national research databases	15,318	80.9% male; 73.2% White; 14.0% Hispanic; 81.0% Cognitively Impaired; Mean age: 9.24 years (4.0)	All participants had at least a "professional diagnosis of ASD"	Parents	CFA (MLE with robust standard errors and Monte Carlo integration); variances fixed at 1, factor covariances freely estimated. Model fit: RMSEA, CFI, TLI, SRMR	1-6 (5)

Factor Analyses of the RBS-R with Non-English-Speaking Populations

In addition to the several studies investigating the factor structure of the original, English version of the RBS-R, four studies have also examined the factor structure of the scale after translation into different languages. Three studies (Georgiades et al., 2010 [EFA only], He et al., 2019 [CFA only]; Kästel et al., 2020 [EFA only]) utilized parent/caregiver ratings to complete their studies, while a third study (Martínez-González & Piqueras, 2017 [EFA and CFA]) used teacher ratings. To date, Martínez-González and Piqueras (2017) are the first to utilize teacher ratings on the RBS-R across all studies, languages, and populations. The following sections include specific details regarding the methods and results of studies utilizing the non-English versions of the RBS-R. See Table 4 (below) for a summary of the studies.

Georgiades et al. (2010)

Georgiades et al. (2010) conducted an EFA with an ASD population to assess the factor structure of the RBS-R for individuals with ASD using the Greek version of the assessment. Authors followed appropriate guidelines for cross-cultural adaptation of measures (Guillemin et al., 1993 as cited in Georgiades et al., 2010) utilizing independently performed forward and backward translations of the instrument. Translators included two native Greek and English speakers with clinical experience in ASD research, which were then assessed independently for accuracy using a consensus committee of Greek and English clinicians, respectively (Georgiades et al., 2010). Authors (Georgiades et al., 2010) highlight the need for this instrument in Greece, as ASD-specific assessments measuring repetitive behavior at that time were scarce.

This study (Georgiades et al., 2010) included a sample of 205 Greek individuals with a diagnosis of ASD, including an age range of two to 48 years old (M = 11.5, SD = 8). Participants were recruited from several regions in the country from a mailing list associated with the Greek

Society for the Protection of Autistic People (Georgiades et al., 2010). All participants were required to provide written evidence (from a physician or clinician) of a clinical ASD diagnosis, though no comprehensive evaluation or confirmation of diagnosis through assessments from associated researchers was required. Parents of the participants included in the study completed all ratings of the RBS-R.

All 43 items from the original RBS-R (Bodfish et al., 2000) were included in a Principal Axis Factoring analysis. The authors performed analyses using a Quartimax Rotation with Kaiser Normalization. Authors (Georgiades et al., 2010) list the Regression Method as the method used to compute factor sores. Factor retention was guided using the scree-plot method, internal consistency of the factors using Cronbach's alpha (Cronbach, 1951), minimum number of factors cross-loadings, and clinical interpretability of the derived factors (Georgiades et al., 2010).

Georgiades et al. (2010) examined one- to six-factor solutions overall, though they selected a two-factor solution based on previously listed criteria. The two retained factors included (I) Compulsive Ritualistic Sameness Restricted Behaviors (CRSRB) and (II) Stereotyped Self-Injurious Behaviors (SSIB). All items from the original RBS-R Compulsive, Ritualistic, Sameness, and Restricted subscales were included in the CRSRB subscale, and all items from the Stereotypy and Self-Injurious Behavior subscales were included in the SSIB subscale. Authors (Georgiades et al., 2010) found that the two-factor solution explained 32.50% of the variance. Additionally, both factors were described as having acceptable internal consistency (CRSB alpha = .92; SSIB alpha = .75; (Georgiades et al., 2010)).

Authors (Georgiades et al., 2010) explain that the purpose of the EFA was not to psychometrically validate the measure but rather to examine the factor structure in the same of Greek individuals with ASD using the RBS-R. However, Georgiades et al. (2010) provide further factor representation of this measure. Furthermore, the two-factor solution provides evidence to support a former school of thought regarding the classification of restricted and repetitive behaviors into “higher-order” and “lower-order” domains (Cuccaro et al., 2003; Georgiades et al., 2010; Papageorgiou et al., 2008; Szatmari et al., 2006). Though the two-factor structure differs from the original six-factor solution (Bodfish et al., 2000) and the proposed five-factor solution (Lam & Aman, 2007), the authors note that past inter-factor correlations (Lam & Aman, 2007) yielded highly correlated factors. Such results suggested a significant amount of shared variance, particularly among the factors collapsed in the 2010 study (i.e., RBS-R Compulsive Behavior, Ritualistic/Sameness Behavior, and Restricted Behavior subscales into the CRSRB subscale; Stereotypic Behavior and Self-Injurious Behavior subscales were included in the SSIB subscale).

Martínez-González and Piqueras (2017)

Martínez-González and Piqueras (2017) conducted an EFA and CFA with an ASD population to assess the factor structure of the RBS-R for individuals with ASD using the Spanish version of the assessment with ratings from teachers and school staff. Authors conducted forward and backward translations of the assessment with supervision from the RBS-R authors (Bodfish et al., 2000) to maintain the fidelity of item and scale content. Authors (Martínez-González & Piqueras, 2017) highlight the need for a structurally sound factor solution for this instrument, with the potential to be used to reach a larger population, expanding beyond English-speaking individuals. Furthermore, the authors underscore the need for the international use of the measures, as repetitive behavior is a core feature of ASD with little measurement or scales to examine symptoms specifically.

Participants included in the study (Martínez-González & Piqueras, 2017) ranged between the ages of three and 63 years (M = 13.00, SD = 9.79), including 233 participants. The sample was majority male (77.70%), White (78.50%), and 15.50% Hispanic. All participants carried an ASD diagnosis from a mental health service and were referred to the study through their local mental health centers. The study included ratings from a range of education staff, including raters from specific special education schools, general education teachers, and daycare center staff.

For the study, authors (Martínez-González & Piqueras, 2017) conducted an EFA using the unweighted least squares (ULS) method for the determination of latent factors underlying the shared variance of items, in addition to a normalized quartimax oblique rotation due to non-normal sampling distributions. The authors also used the direct oblimin extraction method due to the assumed correlation between the factors from previous studies. Additionally, matrix adequacy was tested using KMO and Bartlett's sphericity tests. CFAs of the original six-factor RBS-R (Bodfish et al., 2000) and Lam and Aman's (2007) five-factor version of the RBS-R were also conducted using the Comparative Fit Index (CFI) greater than .95, LISREL Goodness of Fit Index (GFI) greater than .90, RMSEA equal or lower than .08, the Satorra-Bentler chi-square and the Akaike's Information Criteria (AIC). The Turning Point Index (TPI) was used for goodness of fit comparisons between models (Martínez-González & Piqueras, 2017).

For the EFA, all criteria were met to indicate the suitability of the data for performing factor analysis. Following analyses, Martínez-González and Piqueras (2017) presented a six-factor solution, accounting for 43% of the total variance. The proposed six-factor solution includes factors (I) Stereotypic Behavior, (II) Self-Injurious Behavior, (III) Compulsive Behavior, (IV) Ritualistic Behavior, (V) Sameness Behavior, and (VI) Restricted Behavior. The goodness of fit indices provided evidence that the data fit reasonably well in both the five- and six-factor models. However, the six-factor model presented the best fit of all models examined, with all items highly correlating (ranging from .68 to .88) with their associated factors overall.

As mentioned, this study was the first and only to use non-caregiver (e.g., special education and general education staff) ratings for the factor analytic examination of the RBS-R. Though this study utilized a Spanish version of the scale in a population of Spanish-speaking individuals, it should be noted that it is also the only study to support a six-factor model of the RBS-R. While this study utilized a similar factor structure to that of the published RBS-R from Bodfish et al. (2000), this model differs in the assignment of items to factors. Authors do not attribute their varying factor structure to the use of caregiver ratings.

He et al. (2019)

He et al. (2019) examined the appropriateness of the use of the RBS-R with children with ASD in a sample of young children ages three to eight years old in China. First, the study examined the factor structure of the Chinese version of the RBS-R using a CFA approach (He et al., 2019). Second, the study tested the feasibility of using the RBS-R as an assessment tool for ASD identification and diagnosis. In this study, the authors (He et al., 2019) highlight the potential for using the RBS-R, broadly, to study ASD symptoms for assessment, identification, and diagnostic purposes. However, the lack of consensus regarding the factor structure of the instrument, namely in China, is also noted.

For the study, He et al. (2019) sampled children from local early intervention centers, who had a diagnosis of ASD using the DSM-IV-TR (APA, 2000) criteria. Participants included 163 Chinese children with ASD, including 128 males and 34 females. Children's ages ranged from three to eight (M = 5.43) years. The sample of the current study included caregivers as informants on the RBS-R. He et al. (2019) outline translation processes in utilizing the Chinese version of the scale, including guidelines outlined by the International Test Commission Guidelines for Translating and Adapting Tests, Second Edition (International Test Commission, 2017).

Authors (He et al., 2019) do not provide comprehensive data analysis methods in their CFA study. Study methods report examining four competing structural models of the RBS-R, citing Lam and Aman's (2007)'s five-factor model, Mirenda et al. (2010)'s proposed three- and five-factor models, and Martínez-González and Piqueras (2017)'s six-factor model, though it is listed that factor structures for the three- to six-factor models were examined (He et al., 2019). Statistical indices used include a model of chi-square/degrees of freedom, CFI, GFI, RMSEA, SRMR, and AIC (He et al., 2019).

Authors report that the six-factor model yielded GFI, NFI, and CFI results below .70, and considered the five- and six-factor models to be good fits for the data collected. The primary difference between the two models was that Ritualistic and Sameness behaviors comprised one factor in Model V but were separate in Model VI. With this, the authors ultimately chose the five-factor model due to fit statistics and parsimony, like the consensus reached in previous articles (e.g., Lam & Aman, 2007; Mirenda et al., 2010). Though conducted in a different language, this provides further support for the five-factor model of caregiver report of ASD behavior in children with ASD. However, it should be noted that the sample consists of early childhood ranges only.

Kästel et al. (2020)

Kästel et al. (2020) recently examined the factor structure of the RBS-R with an ASD sample using the German version of the instrument. Like similar studies (i.e., Georgiades et al., 2010; Martínez-González & Piqueras, 2017), the authors noted the lack of consistency in the factor structure of the measure overall, but also for non-English speaking populations. Therefore, the study aimed to evaluate the psychometric properties of the instrument overall using a large, diverse sample of German children and adolescents with ASD. In addition to the factor structure of the instrument, item analysis, reliability, and validity data were analyzed and then compared for age and sex differences.

Regarding the factor analytic procedures, Kästel et al. (2020) sampled 948 children overall from four to 17 years (M = 10.55, SD = 4.03). Authors only included a subset of the group for analyses of the RBS-R, due to the nature of the items relating to ASD and associated disorders, though the larger sample included typically developing children, children with ID, and children with “other mental disorders” (Kästel et al., 2020). In the RBS-R analysis group (i.e., ASD group), participants consisted of 504 children with ASD with diagnoses of ASD (43.25%), intellectual disability (ID; 23.81%), and other mental disorders (MD; 32.94%). Participant recruitment took place via flyers, social media, email lists, and school facilities for children with ASD. Participants with ICD-10 diagnoses were included, as well as participants without official diagnoses, but meeting criteria through caregiver-rated SRS and CBCL scales.

Kästel et al. (2020) performed an EFA using OLS estimation and oblimin factor rotation, assuming inter-correlations between factors. All items were assigned to the highest loading factor and factor retention was based on scree plots, Kaiser Criterion for eigenvalues, and factor interpretability. The authors also collected information regarding the internal consistency of the RBS-R total and subscale scores using Cronbach’s alpha estimates with criterion above .90 as excellent, between .80 and .89 as good, between .70 and .79 as fair, and below .69 as unacceptable (Cicchetti, 1994, as cited in Kästel et al., 2020).

Authors cite the calculation of EFA for a two-, three-, four-, five- and six-factor solution. After considering explained variance, the number of items per factor, and item content in each potential factor, the four-factor solution yielded the most evidence for psychometric support. Based on item content, factors were designated as (I) Persistent Behavior, (II) Stereotyped Behavior, (III) Self-Injurious Behavior, and (IV) Compulsive Behavior. The four-factor model was described as explaining 42% of the variance, while the five-factor model explored in this study would have explained 43% of the variance. The authors reported moderately high-to-high correlations between the RBS-R Overall Score and all subscales, excluding the Self-Injurious Behavior subscale. Furthermore, moderate to high correlations are listed between most RBS-R factors (Stereotyped/Persistent Behavior $\propto$ = .64; Compulsive /Persistent Behavior $\propto$ =.80; Compulsive /Stereotyped Behavior $\propto$ = .63; SIB/Stereotyped Behavior = .52; Compulsive/SIB $\propto$ = .43; Persistent/SIB $\propto$ = .41). RBS-R Overall Scores correlations with subscales ranged from $\propto$ = .57 (SIB/Overall Score) to .95 (Persistent Behavior/Overall Score), Median $\propto$ = .85.

Table 4. Summary of Factor Analyses of the Repetitive Behavior Scale-Revised (RBS-R) in Non-English-Speaking Populations

Study and Language	Source	*N*	Sample Demographics	Sample Diagnoses	Rater	Factor Analysis/ Factor Retention Method	Factor Solutions Examined	Factor Solution Names
Georgiades et al. (2010) Greek	Greek Society for the Protection of Autistic People	205	83.9% male; Mean age: 11.5 (8) years; 132 participants under age 12 and 73 over	Required written proof of a clinical ASD diagnosis	Parents	Principal Axis Factoring (PAF) with a quartimax rotation with Kaiser Normalization	1-6; 2 supported	2 factor: I: Compulsive Ritualistic Sameness Restricted Behaviors II: Stereotyped Self-Injurious Behaviors
Martínez-González & Piqueras (2017) Spanish	Local mental health clinics	233	77.70% male; Mean age: 13.00 (9.79) years range from 3-63 years; 78.50% White, 15.50% Hispanic	Required ASD diagnosis from mental health center	Parents	EFA- ULS method, normalized quartimax oblique rotation, direct oblimin extraction CFA- CFI, LISERL GFI, RMSEA, Chi-square, AIC, TPI	1-6; 6 supported	6 factor I: Stereotypic Behavior II: Self-Injurious Behavior III: Compulsive Behavior IV: Ritualistic Behavior V: Sameness Behavior VI: Restricted Behavior

Table 4 (cont'd)

He et al. (2019) Chinese	Early intervention center	163	79.14% male; age ranges 3-8 (*M* = 4.43 years)	Required diagnosis of ASD based on DSM-IV-TR criteria	Parents	CFA – Chi-square/*df*, CFI, GFI, RMSEA, SRMR, AIC	3-6; 5 supported	<u>5 factor</u> I: Stereotypy II: Self-Injurious Behavior III: Compulsive Behavior IV: Ritualistic Sameness Behavior V: Restricted Behavior
Kastel et al. (2020) German	Psychiatric unit at university hospital	948	57.6% male; Mean age: 10.55 (4.03)	ASD 89.4%; ID 90.4%; MD (other mental disorders) 85.4%; TD (typically developin g) 91.8%	Parents	EFA Scree plot, parallel analysis, Bartlett Test of Sphericity with X^2 and Kaiser-Mayer-Olkin	2-6; 4 supported	<u>4 factor</u> I: Persistent Behavior II: Stereotyped Behavior III: Self-Injurious Behavior IV: Compulsive Behavior

Purpose of the Current Study

The purpose of the current study is to examine the psychometric properties of the RBS-R with an ASD sample based on ratings by special education staff. This study will address three specific gaps in the research literature concerning the RBS-R. First, there is a lack of research available to clarify the factor structure of the RBS-R when teachers and other education staff members complete it. This is surprising given the amount of time children with ASD spend in education and/or special education settings. All previously conducted EFAs and CFAs of the English language version of the RBS-R have included samples of individuals with ASD, but none, to the knowledge of this author, has been conducted using ratings completed by special education staff.

However, one study (Martínez-González & Piqueras 2017), though utilizing the Spanish version of the RBS-R in Spain, suggested a six-factor solution for the RBS-R when teachers rated students with ASD. Other available RBS-R studies, all of which relied on caregiver ratings of those with ASD, provided evidence for factor models ranging from two to five factors (Bishop et al., 2013; Georgiades et al., 2010 [Greek]; He et al., 2019 [Chinese]; Hooker et al., 2019; Kästel et al., 2020 [German]; Lam & Aman, 2007; Mirenda et al., 2010; Russell et al., 2019). However, several studies appear to have converged on a five-factor model for this parent/caregiver rater group when rating those with ASD (Bishop et al., 2013; Hooker et al., 2019; Lam & Aman, 2007; Mirenda et al., 2010). Overall, this literature points to limited studies available for other rater types, but the available one suggested a factor structure that is different from those being reported for the other parent/caregiver rater group. Thus, further investigation of the factor structure of the RBS-R is warranted in the context of other rater types—but especially raters in educational settings where children and adolescent students spend a larger part of their time. In addition, educational professionals need progress monitoring tools for their students with these types of needs. With sufficient psychometric evidence to support it in this role, the RBS-R could be a very useful tool for this purpose.

Second, there is a lack of sufficient and adequate research studies to support and achieve consensus regarding the most appropriate factor structure for the RBS-R in samples of individuals with ASD. Though several studies have examined the factor structure of the RBS-R, there still exists uncertainty regarding the most appropriate, generalizable factor structure. As a widely used instrument in the field of RRBIs and ASD, the RBS-R needs a supported factor structure to inform and support a viable and meaningful subtest structure that can then be subject to other types of validation (e.g., convergent, and divergent relationships with other measures of the same or different theoretical construct—across a consistent set of subscales).

Third, no current published study has performed a CFA on the RBS-R directly comparing all the available factor models generated with ASD samples—even if restricted to just the English language RBS-R and parent/caregiver raters. The current study will address RBS-R-specific needs consistent with relevant guidelines presented in the SEPT (2014) standards concerning assessment tool validity and test design and development, as discussed earlier in the review of the literature. Concerning validity, Standards 1.1, 1.2, 1.3, and 1.4 underscore the principle that tests must be revised and/or further validated when used with new populations or for purposes other than those initially intended and covered by extant validity studies (SEPT, 2014). The present research literature has not established satisfactory evidence for the use of the RBS-R based on special education staff ratings, though it is used commonly in practice based on the factor structure proposed initially with caregiver ratings. Regarding test design and development, Standards 1.13, 4.6, and 4.24 provide a more specific set of guidelines on the test development process. The former two standards highlight the importance of developing tests further as new data arises. Standard 4.24 indicates that test specifications are often needed for revisions to the instrument based on the presence of new data. Thus, exploration and validation of the RBS-R factor structure for those with ASD rated by special education staff is a clear area of need in the current research literature and essential to inform practice.

Research Questions

Questions one through five, described below, will be investigated using exploratory factor analytic techniques. Questions six and seven will be investigated using confirmatory factor analytic techniques. Research questions one through five will be included in the method subsection for study one and research questions six and seven will be included in the method subsection for study two.

Research Question 1. When special education staff rate students with ASD, how many possible or likely interpretable RBS-R factors are available for retention consideration?

Research Question 2. When special education staff rate students with ASD, how many factors should be retained to derive the most interpretable factor solution for the RBS-R?

Hypothesis 2a. It is predicted that the most interpretable EFA solution will consist of multiple factors (i.e., $\geq$ 2 factors).

Hypothesis 2b. It is predicted that, within the most interpretable EFA solution, a Self-Injurious Behavior factor will be present.

Hypothesis 2c. It is predicted that, within the most interpretable EFA solution, a Compulsive Behavior factor will be present.

Research Question 3. When special education staff rate students with ASD, does the most interpretable factor structure yield substantive correlations among at least some of the factors of the RBS-R?

Research Question 4. When special education staff rate students with ASD, to what extent does the six-factor RBS-R solution from the present EFA (regardless of whether this was the retained solution) correspond to the six-subscale structure originally proposed by Bodfish et al. (2000)?

Research Question 5. When special education staff rate students with ASD, to what extent does the five-factor RBS-R solution from the present EFA (regardless of whether it was the retained solution) correspond to the five-factor solution reported by Lam and Aman (2007)?

Research Question 6. When special education staff rate students with ASD, does the model generated by the EFA of the RBS-R in the exploratory sample fit the inter-item covariance matrix of the CFA validation sample?

Hypothesis 6. It is predicted that the most interpretable model generated by the EFA in sample one, will adequately fit the inter-item covariance matrix of the CFA validation sample.

Research Question 7. In the confirmatory sample, where special education staff rate students with ASD, how does the factor solution generated by the EFA of the RBS-R compare in terms of absolute and relative fit to the previous RBS-R factor models found in ASD samples (e.g., Hooker at al., 2019; Martinez-Gonzalez & Piqueras, 2017; Mirenda et al., 2010)?

Hypothesis 7. It is predicted that the most interpretable model generated by the EFA in sample one, will demonstrate better relative fit to the inter-item covariance matrix of the CFA sample when compared to previously reported RBS-R factor models derived from ASD samples.

CHAPTER THREE: METHOD

The present dissertation consists of two interrelated studies. The first study involves an EFA intended to address research questions one through five. In the second study, a CFA informed by the results of the EFA from study one answers research questions six and seven. Research design and procedures used for data collection and analyses, study hypotheses, and associated methods are discussed below.

Research Design

The focus of both studies is on instrument validation with an emphasis on exploring and confirming the internal structure validity of the RBS-R, its factor structure, and model fit, within a sample of students with ASD who were rated by special education staff. The study design falls within the scope of an observational, correlational, and cross-sectional approach (Kazdin, 2017). These classifications are used, as the study involves measuring variables at a single time point in a large group of participants and the in-depth assessment of relations among those variables. In terms of purpose, the in-depth assessment of relations among the RBS-R items involves the use of multivariate statistical techniques to examine the measure's latent structure and better understand the shared meanings underlying the inter-item relations. More specifically, factor-analytic techniques used for the determination of the most interpretable and substantive number of potential underlying constructs that may explain shared variability among the RBS-R items, assess the fit of the EFA-derived model in a confirmatory sample and compare the fit of the EFA-derived factor model to other RBS-R factor models available within the research literature.

Extant Data Collection

Participant assessment data for studies one and two were obtained from an existing raw data archive based in a large special education agency in Western New York State that serves students with ASD and other significant developmental disabilities. Extensive rating scale data were originally collected for purposes of progress monitoring and annual program evaluation. Participants included in both study one and study two were current students at the agency at the time of data collection. Though most students had data available from multiple time points, only data from a single time point for each participant were used within each factor analytic sample. Thus, each participant contributed items from no more than one RBS-R administration for the given factor analysis.

Raters

All the original data collection occurred annually for program evaluation purposes between the years 2009 and 2020. The existing dataset, resulting from these program evaluation activities, includes participant ratings completed by special education staff members. The staff members were individuals familiar with each participant in the special education classroom environment. Raters consisted of special education teachers, teaching assistants, speech/language pathologists, physical therapists, occupational therapists, behavior technicians, individual student aides, whole-classroom aides, and trained volunteer assistants employed by the special education agency. Agency psychologists made rater assignments for the staff. A large number of staff available within each classroom allowed, in most cases, for staff members to rate only one student each. This one-to-one rater-student correspondence was intentionally prioritized to minimize the potential influence of multiple ratings being nested within individual raters. However, whenever possible, the rater who knew the student best was selected to perform the rating. Across the sample, rater familiarity with each student ranged from six weeks to twenty-eight months of regular interaction.

Procedures

Data collection procedures were developed and carried out by the special education agency, as part of its annual program evaluation process. On an annual basis, each rater was allocated a packet of five rating measures to be completed for their assigned student. All included measures were counter-balanced at random within each packet and staff members were instructed to complete all instruments in the given order. Following packet completion, a program evaluation staff (usually an agency psychologist) reviewed all protocols for completeness and for possible errors (e.g., multiple responses to the same item, omitted responses to items, etc.). In the case of a missing or incorrectly rated item response, staff reviewers contacted the original rater to correct the identified errors. After completion of all measures, two program evaluation staff independently scored each measure. When necessary, a third program evaluation staff independently scored the measure to resolve any discrepancies between the primary scoring staff.

The director of program evaluation at the agency assigned each student a unique identification code. Only these codes were used on packets and rating measures. No other identifying information was included. The director of program evaluation was the only individual with access to the code key that linked identifying information to each code. The investigator for the present study does not have access to any individual identifying information beyond the case identification code, as all rating measures were otherwise de-identified at the time of data collection at the agency.

Inclusion/Exclusion Criteria

Determination of participant suitability for study inclusion followed a three-stage screening process. First, all participants' ages must have been between the ages of three and 21 years old—spanning preschool to the end of high school within special education. Second, all participants were required to have a clinical diagnosis of autistic disorder or PDD-NOS based on the DSM-IV-TR (APA, 2000) criteria or an ASD diagnosis based on DSM-5 (APA, 2013) criteria. Participants with ASD were not required to have but were included if they did have a co-occurring diagnosis of ID or other co-morbid DSM-5 diagnoses (e.g., attention-deficit/hyperactivity disorder [ADHD], language disorder, etc.) that are frequently observed in the context of ASD. A licensed professional (e.g., licensed psychologist, licensed medical professional) determined all diagnoses. The inclusion of individuals with an ASD special education eligibility designation, as determined by the participant's school-based special education committee was also allowed. In cases where the special education designation of ASD was the primary diagnostic source, an agency psychologist confirmed that DSM-5 criteria for ASD were also met. Third, all participants must have been, at one time, active participants in special education classrooms for students considered to have a substantial functional impairment (e.g., individuals with significant delays in cognitive, social, or communication domains—sufficient to warrant center-based special educational programming).

Due to substantial variability in participant ages, cognitive development, communication skills, and individual behavioral challenges, agency psychologists were not able to utilize a uniform measure of cognitive ability across all participants. As a result, estimates of general cognitive ability were derived from a variety of cognitive measures -- reflecting the different testing needs of students who varied on these dimensions. The global score from each of these instruments was converted into a common deviation quotient metric--with a normative mean of 100 and a standard deviation of 15. This was done so that overall cognitive scores could be pooled for purposes of describing the sample. Measures used to assess cognitive ability included:

Bayley Scale of Infant Development (Bayley, 1969); Bayley Scales of Infant Development, Second Edition (Bayley, 1993); Bayley Scales of Infant Development ,Third Edition (Bayley, 2006); Stanford-Binet Intelligence Test, Fourth Edition (SB4; Thorndike et al., 1986); Stanford-Binet Intelligence Test, Fifth Edition (SB-5; Roid, 2003); Comprehensive Test of Nonverbal Intelligence (CTONI; Hammill et al., 1996); Cognitive Assessment System, Second Edition (CAS2; Naglieri et al., 2014); Differential Ability Scales (DAS; Elliott, 1990); Differential Ability Scales, Second Edition (DAS-II; Elliott, 2007); Kaufman Assessment Battery for Children (KABC; Kaufman & Kaufman, 1983); Kaufman Brief Intelligence Test (KBIT; Kaufman & Kaufman, 1990); Learning Accomplishment Profile - Diagnostic Standardized Assessment (LAP-D; Nehring.et al., 1992); McCarthy Scales of Children's Abilities (MSCA; McCarthy, 1972); Universal Nonverbal Intelligence Test (UNIT; Bracken & McCallum, 1998); Wechsler Abbreviated Scale of Intelligence (WISC; Wechsler, 1999); Wechsler Abbreviated Scale of Intelligence, Second Edition (WASI-II; Wechsler, 2011); Wechsler Adult Intelligence Scale, Third Edition (WAIS-III; Wechsler, 1997); Wechsler Intelligence Scale for Children, Revised (WISC-R; Wechsler, 1974); Wechsler Intelligence Scale for Children, Third Edition (WISC-III; Wechsler, 1991); Wechsler Preschool and Primary Scale of Intelligence, Revised (WPPSI-R; Wechsler, 1989); Wechsler Preschool and Primary Scale of Intelligence, Third Edition (WPSSI-III; Wechsler, 2002); and Wechsler Preschool and Primary Scale of Intelligence, Fourth Edition (WPPSI-IV; Wechsler, 2012).

Measure

The RBS-R (Bodfish et al., 2000) is a rating scale and assessment tool used to assess the variety and perceived severity of restricted, repetitive behaviors present in individuals with ASD. The current RBS-R is the second iteration of the original RBS (Bodfish et al., 2000), which was initially developed as a collection of four “checklists” (i.e., Stereotypy, Self-injury, Compulsion, and Tic checklists) for collecting information about the topography and severity of repetitive behaviors observed in everyone. There is no manual for the RBS or RBS-R that could provide more specific administration guidance and summarize psychometric evidence. However, several studies have examined the measurement characteristics of this assessment tool (Bishop et al., 2013; Hooker et al., 2019; Lam & Aman, 2007; Mirenda et al., 2010; Rojahn et al., 2013; Russell et al., 2019; Schertz et al., 2016) and their relevant psychometric findings are presented in this section.

In its present form, the RBS-R is a 43-item rating scale developed to assess six broad domains of repetitive behavior. The subscales on the RBS-R differ from the scale’s first iteration in that the RBS subscales were divided into four domains: Stereotypic Behavior, Self-Injurious Behavior, Compulsions, and Tics, while the RBS-R is comprised of six subscales. These six subscales are Stereotyped Behaviors (6 items), Self-Injurious Behavior (8 items), Compulsive Behavior (8 items), Ritualistic Behavior (6 items), Sameness Behavior (11 items), and Restricted Behavior (4 items; Bodfish et al., 2000). Initially, the RBS-R items were conceptually placed into each domain based on the authors’ clinical experience (Bodfish et al., 2002, reported in Lam & Aman, 2007). Scale authors (Bodfish et al., 2002) then conducted a principal components analysis (PCA) using a small sample (N = 124). The authors retained a six-component solution (Bodfish et al., 2002, reported in Lam, 2004) that included an unusual number of problematic cross-loadings. The six-component model roughly corresponded to Bodfish et al.’s (2002) conceptual placement of items into six factors. However, there were several discrepancies in item assignments across the identified components, as some items loaded similarly or higher on components (factors) to which they were not originally assigned (Bodfish et al., 2002 as cited in Lam, 2004).

The rater responds to each RBS-R item using a four-point Likert scale ranging from zero to three, which conveys the presence and severity of each behavior. Scale response anchors are *behavior does not occur* = 0, *behavior occurs and is a mild problem* = 1, *behavior occurs and is a moderate problem* = 2, and *behavior occurs and is a severe problem* = 3. At the subscale level, informants are asked to provide information about the frequency and interference of the behavior that is the focus of the subscale. Informants are further asked about the level of distress caused by the interruption of that behavior. These additional ratings are made by marking a visual analog rating scale (VARS) for each dimension with dimension-relevant anchor labels at each end (i.e., "never" to "constant" for frequency, "not at all" to "severe interference" for interference, and "not at all" to "severe" for distress). After completing all other items, informants are also asked to provide a global rating of how much of a problem the individual's repetitive behaviors are overall using a VARS method. The global score can range numerically from one (i.e., verbal anchor = "not a problem at all") to 100 (i.e., verbal anchor = "as bad as you can imagine"). The use of these additional VARS ratings is not typically included in the literature examining the psychometric properties of the RBS-R.

Item raw scores are summed for each RBS-R subscale (i.e., total raw subscale scores for Stereotyped Behavior, Self-injurious Behavior, Compulsive Behavior, Ritualistic Behavior, Sameness Behavior, and Restricted Behavior) and a total RBS-R score (i.e., the overall score for the instrument). However, the RBS-R has not been standardized using a normative sample. Thus, no norm-referenced standard scores are available and no cut scores or scoring interpretation guidelines are provided. The authors also do not provide information about potentially important informant prerequisites, requirements, or characteristics (e.g., relationship to the individual, how long the informant has known the individual, locations the informant is familiar with the individual, etc.) needed to qualify as an appropriate person or completing the rating scale. In most research studies involving the RBS-R, raters have been parents/caregivers of the rated person with ASD.

RBS-R Reliability

Generally, the RBS-R has demonstrated adequate to good levels of reliability for research purposes across studies but results for subscales have shown much more variability in terms of meeting higher standards for clinical use. Without an official RBS-R manual from the rating scale authors, all psychometric data must be extracted from published studies. In consideration of appropriate internal consistency benchmarks, reliability coefficients should be at least .70 to meet the minimum standard for research purposes, and at least .90 minimum for individual decision-making (Nunnally, 1978). Additionally, Murphy and Davidshofer (1988) suggested standards where estimates $\leq$ .59 are considered very poor measures of reliability, estimates ranging from .60 to .69 as low to poor measures of reliability, ranging from .70 to .79 as moderate measures of reliability, ranging from .80 to .89 as good reliability, and $\geq$ .90 as excellent measures of reliability. Extending these interpretive guidelines, or benchmarks, for reliability, Salvia et al. (2017) suggested that measures used for high-stakes situations (e.g., special education decision-making) should have reliability coefficients of at least .90, measures used for screening purposes should have a reliability of at least .80, and measures used for frequent progress monitoring should have reliability coefficients of at least.70.

Internal Consistency Reliability. Consistent with the above, internal consistency reliability estimates for the instrument’s total scores and subscales have generally ranged from moderate (≥ .70) to excellent (≥ .90) for research purposes (Murphy & Davidshofer, 1988). As indicated previously (see Chapter 2), several researchers have adopted the five-factor model of the RBS-R (e.g., Esbensen et al., 2009; Lam & Aman, 2007; Mirenda et al., 2010; Sturm et al., 2022). Therefore, several independent studies have reported on the reliability and validity of the five subscales reflecting the five-factor model instead of subscales based on the original six-factor model.

In an independent study using ratings from caregivers of children and adolescents with ASD, Lam and Aman (2007) reported the internal consistency of the five-subscale model of the RBS-R, ranging from .78 (Restricted Interests) to .91 (Ritualistic/ Sameness Behaviors), with a median estimate of .84 (Self-Injurious Behavior). There were no data reported for the internal consistency of a Total Score. Esbensen et al. (2009) reported an internal consistency estimate of .93 for the RBS-R Total Score, while RBS-R subscales (based on Lam & Aman's (2007) five-factor model] ranged from .74 (Restricted Interests) to .89 (Ritualistic/Sameness), with median reliability of .77 (Self-Injurious Behavior). Internal consistency estimates were similar when examined for male and female participants, individuals with ASD only, and individuals with ASD and ID combined (Esbensen et al., 2009). Further, Mirenda et al. (2010) reported on the internal consistency of both the five- and six-factor models of the RBS-R in a sample of young children with ASD rated by their caregivers. For the six-factor model, internal consistency estimates ranged from .71 (Ritualistic Behavior) to .88 (Sameness Behavior), with a median estimate of .76. For the Lam and Aman (2007) five-factor model, internal consistency estimates ranged from .72 (Compulsive Behavior) to .90 (Ritualistic Behavior), with a median estimate of .73 (Stereotypic Behavior). Miranda et al. Did not report an internal consistency estimate for an

Overall RBS-R Score. Sturm et al. (2022) reported internal consistency estimates ranging from .66 to .88, on Factors V and III, respectively, though no specific factor names were provided.

Rojahn et al. (2013) report on the internal consistency of the RBS-R utilizing the original six-factor model with ratings made by caregivers of young children (ages four to 48 months, *M* = 27.4 months, *SD* = 10.1) with ASD. Internal consistency estimates using Cronbach's alpha ranged from .30 (Restricted Behavior) to .75 (Stereotyped Behavior), with a median estimate of .70 and an Overall Score estimate of .89. The relatively lower internal consistency estimates for the subscales in this last study may have been due to the very young age of the children being rated (e.g., if some items were not relevant for the youngest children and responses resulted in range restriction). = 10.1) with ASD. Internal consistency estimates using Cronbach's alpha ranged from .30 (Restricted Behavior) to .75 (Stereotyped Behavior), with a median estimate of .70 and an Overall Score estimate of .89. Similarly, Schertz et al. (2016) reported acceptable Cronbach's alpha estimates ranging from .74 (Self-Injurious Behavior) to .92 (Stereotyped Behavior), with a median reliability estimate of .79 using the six-factor solution of the RBS-R.

Interrater Reliability. Broadly, interrater reliability estimates have ranged from low ($\geq$ .60) to good ($\geq$ .80) for research purposes (Murphy & Davidshofer, 1988). RBS-R authors reported that the original RBS indicated good Overall Score interrater reliability (.88), though more specific information regarding the interrater reliability of the four checklists was not provided. Additional reliability estimates were reported by Bodfish and Lewis (2002) and by Lam and Aman (2007) for the second iteration of the scale, the RBS-R. Based on the original six-factor subscale model, subscale interrater reliability ranged from .55 (Sameness Behavior) to .78 (Self-Injurious Behavior), with no median or Overall Score estimates reported. Lam and Aman (2007) reported interrater reliability for their five-factor solution of subscales ranging from .57 (Compulsive Behavior) to .73 (Stereotypic Behavior), reporting median reliability of .67. For the Overall Score, Lam and Aman (2007) report an interrater reliability estimate of .70.

Test-Retest Reliability. Additional studies have reported test-retest reliability estimates for the measure, though evidence, in general, is lacking in the literature. RBS-R authors reported that the original RBS yielded a test-retest reliability estimate of .71 (Bodfish et al., 2000). Test-retest reliability data for the RBS-R subscales yielded a wide range of values, ranging from .52 (Ritualistic Behavior) to .96 (Restricted Interests; Bodfish & Lewis, 2002; Lam & Aman, 2007). Similar to previously reported studies, no information regarding test-retest reliability estimates for an RBS-R Total Score was reported in either study.

RBS-R Validity

Regarding internal structure, several factor analytic studies of the original English version of RBS-R, wherein parents/caregivers rated samples of individuals with ASD, have been reported in the literature. However, a range of factor solutions have been retained across these studies and a clear consensus regarding the number of factors present has not yet been reached. In the United States, three EFA studies (Bishop et al., 2013; Lam & Aman, 2007; Russell et al., 2019) and one CFA study (Hooker et al., 2019) were conducted with the English version of the instrument, involving caregiver ratings of participants with ASD. One CFA study (Mirenda et al. 2010) of the English version of the RBS-R conducted in Canada also involved caregiver ratings of participants with ASD. Across studies, various factor selection criteria have suggested examining factor solutions ranging from two to six factors (EFAs: Bishop et al., 2013; Lam & Aman, 2007; Russell et al., 2019; and CFAs: Hooker et al., 2019; Mirenda et al., 2010). However, among these studies, only four- (Russell et al., 2019) and five-factor (Bishop et al., 2013; Hooker et al., 2019; Lam & Aman, 2007; Mirenda et al., 2010) solutions have been retained as most interpretable and meaningful or judged to be the best fit. When foreign language translations of the RBS-R are considered, various factor analytic studies involving parent/ caregiver ratings of children and/or adolescents with ASD have retained two- (Georgiades et al., 2010 [Greek RBS-R]), four- (Kästel et al., 2020 [German RBS-R]), and five-factor (He et al., 2019 [Chinese RBS-R]) solutions as most interpretable. In addition, within this foreign translation literature, Martínez-González and Piqueras (2017)'s EFA and CFA of the Spanish translation of the RBS-R used professional staff and teacher ratings of students with ASD. These authors ultimately retained a six-factor solution as the most interpretable. This study is noteworthy for being the only available factor analytic study of the RBS-R that used professional staff and teacher (not parent/caregiver) ratings of students with ASD. It is also the only study, outside of the limited early PCA by Bodfish (2002, as cited in Lam, 2004) to retain a six-factor solution. (Details of each EFA and CFA study are included in Chapter 2.)

Some evidence of concurrent validity has been established amongst subscales of the RBS-R and subscales on two existing measures, the Behavior Problems Inventory (BP-01; Rojahn et al., 2013) and the ABC (Aman, 1985), though wider concurrent validity evidence for the RBS-R subscales is still underdeveloped. (Note that a consensus regarding the RBS-R factor structure would help in this area of validity, as it would more clearly establish the number of subtests and their hypothesized constructs for purposes of concurrent validity efforts.) Generally, the comparisons including the RBS-R and other measures yielded relatively lower-than-expected validity coefficients that ranged between fair and excellent. When compared to the BP-01 Self-Injurious Behavior subscale, the RBS-R Self-Injurious Behavior subscale was significantly correlated (.54). This pattern exists for the two Stereotypy subscales on both measures as well (.41). When compared to the ABC, there was a significant correlation between the ABC and

RBS-R Stereotypy subscales (.55) and the ABC Irritability subscale (which includes three items relating to self-injury) and the RBS-R Self-Injurious Behavior subscale (.38) in one study (Rojahn et al., 2013).

Mirenda et al. (2010) also provided some support for the convergent validity of their three-factor and five-factor models of the RBS-R through associations between factor-based subscales and scores across several subscales of the CBCL (Achenbach & Rescorla, 2000). For both models, the strongest correlations were with the CBCL total raw scores (i.e., not norm-referenced scores) and subscale raw scores. All factors in the three- and five-factor models (except Self-Injurious Behavior in Factor V) were correlated moderately with the ADI-R repetitive behaviors total score (Mirenda et al., 2010). See Table 5 (below) for specific Spearman's rank-order correlation coefficients for Models III and IV. In their CFA study, Hooker et al. (2019) provided estimates of convergent validity between their five-factor model of the RBS-R and the ADOS-2 Restricted and Repetitive Behavior (RRB) Subscale. Associations with the ADOS-2 RRB domain score yielded statistically significant correlations between three of five factors of their RBS-R solution using summed factor scores (SIB subscale $Z = -2.17$**, Compulsive Behavior subscale $Z = -2.17$**, Restricted/Sameness Behavior subscale $Z = -1.98$**).

Table 5. Spearman's Rank-Order Correlation Coefficients for Models III and V Included in Mirenda et al. (2010)

	RBS-R Factor							
	Model III: CRSB	Model III: SIB	Model III: RSB	Model V: STEREO	Model V: SIB	Model V: COMP	Model V: RIT/ SAME	Model V: REST
CBCL Internalizing Problems Raw Score	.629**	.384**	.550**	.455**	.384**	.476**	.619**	.494**

Table 5 (cont'd)

CBCL Externalizing Problems Raw Score	.531**	.437**	.517**	.452**	.437**	.394**	.531**	.444**
CBCL Total Raw Score	.648**	.497**	.608**	.528**	.497**	.491**	.639**	.519**
ADI-R Repetitive Behaviors Domain Total	.377**	.129*	.374**	.297**	.129**	.330**	.362**	.351**

Note: RBS-R Repetitive Behavior Scale – Revised, *CBCL* Child Behavior Checklist, *ADI-R* Autism Diagnostic Interview – Revised

* $p < .05$ level (2-tailed), ** $p < .01$ level (2-tailed)

Generally, evidence of discriminant validity regarding the RBS-R has not been examined thoroughly at this time. When compared to the ABC as a discriminant measure of validity, there was one weak, non-significant correlation (.09) between scores of theoretically independent subscales on the two measures (ABC *Excessive Speech* subscale and RBS-R *Self-Injurious Behavior*s subscale). When compared to the BP-01, there were no examples of non-significant, weak correlations between scores of theoretically independent subscales (Rojahn et al., 2013). The authors reported that identifying evidence for discriminant validity for the included measures was difficult due to the complex relationships between aberrant behavior, psychopathology, and repetitive behavior (Rojahn et al., 2013).

Study One: EFA

Research Questions, Rationales, and Hypotheses

Questions one through five were examined via exploratory factor analytic techniques in study one, while questions six and seven will be examined via confirmatory factor analytic techniques in study two. For study one, based on RBS-R ratings made by special education staff among a sample of individuals with ASD:

Research Question 1. When special education staff rate students with ASD, how many possible or likely interpretable RBS-R factors are available for retention consideration?

Research Rationale. Three studies involving the original English version of the RBS-R have examined the factor structure of the RBS-R with a sample of individuals with ASD (Bishop et al., 2013; Lam & Aman, 2007; Russell et al., 2019). (All involved only parent/caregiver ratings.) Among these factor analyses, factor selection criteria (e.g., Kaiser criterion, scree plot, etc.) suggested between two and six factors was likely the range in which the most interpretable solution would be found. Ultimately, either a four- or five-factor solution was chosen as the most appropriate model in each study. Bishop et al. (2013) report that results supported a five-factor solution based on cutoffs (0.30 or higher) for including items on a factor and based on previous studies (e.g., Lam & Aman, 2007). In the EFA by Lam and Aman (2007), factor solutions consisting of between two and six factors were examined for interpretability (e.g., scree plot method, Cattell, 1966; eigenvalues above 1.0; goodness of fit as estimated by RMSEA, Browne & Cudeck, 1992; and clinical interpretability). Ultimately, a five-factor solution was retained as the most interpretable and meaningful. Finally, Russell et al., (2019) examined the interpretability of factor solutions consisting of between two and five factors in their EFA. The authors retained a four-factor solution as the most interpretable.

Additionally, several studies have examined the factor structure of the RBS-R in foreign languages using EFAs. Of note, three studies (Georgiades et al., 2010; Kästel et al., 2020; Martínez-González & Piqueras, 2017) utilized Greek-, German-, and Spanish translations of the RBS-R, respectively; however, authors reported that the translations did not change the meaning or application of item content. Results of a study (Georgiades et al. (2010) examining the Greek version of the RBS-R indicated that the two-factor solution was best supported by analyses, based on criteria considering a scree-plot, internal consistency of factors, minimum number of factors cross-loadings, and clinical interpretability of the derived factors. The two-factor solution was reported to explain 32.50% of the variance. Kästel et al. (2020) retained a four-factor solution for the German RBS-R. The criterion for retention included a scree plot and parallel analysis. Following their EFA of the Spanish version of the RBS-R, Martínez-González and Piqueras (2017) retained a five- and six-factor solution for the instrument. Authors (Martínez-González & Piqueras, 2017) also conducted a CFA following the EFA, reporting the six-factor solution as the best fit.

In the present study, the EFA was conducted using Principal Axis Factoring (PAF), and initial factor selection criteria included the traditional Guttman-Kaiser criterion (Guttman, 1954; Kaiser, 1960) and scree test (Cattell, 1966), as well as the more sophisticated and accurate techniques of parallel analysis (Horn, 1965) and Velicer's minimum average partial test (MAP; Velicer, 1976). After the criteria were met, a range of factor solutions will be explored. Specifically, solutions including the range of factors meeting the criteria, plus or minus two factors will be examined for interpretability.

Research Question 2. When special education staff rate students with ASD, how many factors should be retained to derive the most interpretable factor solution for the RBS-R?

Research Rationale and Hypothesis 2a, 2b, 2c. Previous factor analyses examining the factor structure of the RBS-R using a sample of individuals with ASD have resulted in the retention of two-, four-, five-, and six-factor solutions. Past analyses include studies using the English RBS-R EFAs: Bishop et al. (2013), Lam and Aman (2007), and Russell et al. (2019).

EFAs in other cultures using the RBS-R in translation include the Greek RBS-R from Georgiades et al. (2010), German RBS-R from Kästel et al. (2020), and Spanish RBS-R from Martínez-González and Piqueras (2017). Georgiades et al. (2010) selected a broader two-factor solution (Compulsive Ritualistic Sameness Restricted Behaviors and Stereotyped Self-Injurious Behaviors). Russell et al. (2019) selected a four-factor solution (Rituals/Sameness, Compulsive Behaviors, Stereotypic Behaviors, and Self-Injurious Behaviors). Kästel et al. (2020) also selected a four-factor solution (Persistent Behavior, Stereotyped Behavior, Self-Injurious Behavior, and Compulsive Behavior). Lam and Aman (2007) chose a five-factor solution (Rituals/Sameness, Self-Injurious Behavior, Stereotypic Behavior, Compulsive Behavior, and Restricted Interests). Likewise, Bishop et al. (2013) found evidence to support a five-factor solution (Sensory-Motor Behavior, Restricted Interests, Self-Injurious Behavior, Compulsive Behavior, and Ritualistic Behavior) consistent with Lam and Aman's (2007)'s study, renaming the previous "Stereotypic Behavior" factor, Sensory-Motor Behavior – both factors are identical at the item-level. Like the originally proposed six-subscale structure of the RBS-R, Martínez-González and Piqueras (2017) selected a six-factor solution (Stereotypic Behavior, Self-Injurious Behavior, Compulsive Behavior, Ritualistic Behavior, Sameness Behavior, and Restricted Behavior) as most interpretable.

Across the available factor-analytic studies of the RBS-R (Bodfish et al., 2000), factors reflecting compulsive behavior and self-injurious behavior have been retained consistently. All studies have also included aspects of both ritualistic behavior and sameness behavior, consistent with the original RBS-R (Bodfish et al., 2000). However, all studies supporting a four- and five-factor solution combined these two aspects or expressions of repetitive behavior together, instead of separating the two into distinct factors. After considering the existing factor analyses studies, three hypotheses were made: a) at least two factors would be retained, b) a Self-Injurious Behavior factor would emerge, and c) a Compulsive Behavior factor would emerge. All three hypotheses were tested by examining the pattern and structure matrices, resulting from oblique direct oblimin rotation (Jennrich & Sampsons, 1966) of factors extracted via PAF, for interpretability of factors across the range of potential factor solutions.

Research Question 3. When special education staff rate students with ASD, does the most interpretable factor structure yield substantive correlations among at least some of the factors of the RBS-R?

Research Rationale. The interpretive phase of an EFA begins with assessing the possible correlations among the factors (Fabrigar et al., 1999; Preacher & MacCallum, 2003). By examining the correlations among factors, a better understanding of the strength of the relations between underlying constructs or latent variables emerges. Some constructs will or should be more correlated or less correlated depending on the nature or theoretical understanding of each construct (Kline, 1994). Lam and Aman (2007) reported inter-factor correlations ranging from .14 (Restricted Interests and Self-Injurious Behavior) to .48 (Self-Injurious Behavior and Stereotypic Behavior) within their retained five-factor solution. Georgiades et al. (2010) reported an inter-factor correlation of .05 in their two-factor solution (Compulsive Ritualistic Sameness Restricted Behaviors and Stereotyped Self-Injurious Behaviors) for the Greek RBS-R. Bishop et al. (2013) reported inter-factor correlations ranging from .25 (Restricted Interests and Self-Injurious Behavior) to .67 (Ritualistic/Sameness Behavior and Compulsive Behavior) in their five-factor solution of the RBS-R. Russell et al. (2019) reported inter-factor correlations ranging from .28 (Compulsive Behavior and Self-Injurious Behavior) to .58 (Rituals/Sameness and Compulsive Behavior) for their four-factor solution. Kästel et al. (2020) reported inter-factor correlations ranging from .41 (Persistent Behavior and Self-Injurious Behavior) to .80 (Persistent Behavior and Compulsive Behavior) for the German RBS-R. Additionally, Martínez-González and Piqueras (2017) report inter-factor correlations ranges from .30 (Ritualistic and Self-Injurious) to .68 (Sameness and Compulsive) for the Spanish RBS-R. Results of the previously published EFA studies examining the factor structure of the RBS-R (in both English and foreign language versions) provide a range of expected correlations amongst the proposed factors. Therefore, it was believed that there would be substantial correlations among at least some factors in the most interpretable solution.

Research Question 4. When special education staff rate students with ASD, to what extent does the six-factor RBS-R solution from the present EFA (regardless of whether this was the retained solution) correspond to the six-subscale structure originally proposed by Bodfish et al. (2000)?

Research Question 5. When special education staff rate students with ASD, to what extent does the five-factor RBS-R solution from the present EFA (regardless of whether it was the retained solution) correspond to the five-factor solution reported by Lam and Aman (2007)?

Research Rationales. To establish an open exploratory approach to the analysis of the RBS-R in this study, while limiting any bias from preconceptions, all possible factor solutions suggested by factor selection criteria were meticulously analyzed and interpreted independently by several ASD researchers.

However, there are two different factor models for the RBS-R (i.e., the original six-factor model implied by the six-subscale interpretive framework proposed by the RBS-R authors and the five-factor model retained by Lam & Aman [2007]) that hold some important status within the present research literature. The six-subscale model by Bodfish et al. (2002; cited in Lam, 2004), is important because its structure was used to frame and organize the items on the RBS-R itself. At the same time, the five-factor model from Lam and Aman (2007) has been replicated to some extent in two other studies (see Bishop et al., 2013; Hooker et al., 2019) and has become more widely accepted as an alternative basis for scoring the RBS-R relative to other factor models reported in the literature.

The question arises, regardless of the most interpretable solution retained in any EFA, have other known factor solutions emerged among the various factor solutions extracted as part of the analysis. Using the Lam and Aman (2007) five-factor model as an example, listing different possible outcomes of this type of analysis illustrates the potential value of asking such questions for the larger RBS-R research literature. First, a researcher may end up retaining a five-factor solution as the most interpretable and meaningful from among those available. In this instance, this five-factor solution may align well in terms of factor interpretations and item content for each factor with the Lam and Aman model. This would be a direct replication of the five-factor solution reported by Lam and Aman, which could move the literature closer to consensus on the most appropriate factor structure. Second, suppose that the researcher retains a seven-factor model as most interpretable and meaningful, but when a five-factor solution is extracted, it is very similar to what Lam and Aman found. This would suggest that despite the replication of a previously reported factor structure, the present study suggested that there may be an even better factor model available—perhaps one that is more complex and nuanced than the five-factor model. Third, suppose a seven-factor model is retained and when the five-factor model is examined, it does not line up well with the previously reported five-factor model from the literature. In this case, the five-factor model was not replicated at all and a different factor solution consisting of more factors was retained as the most interpretable.

This complete lack of replication would be noteworthy and contribute to the need to better understand why different samples are producing discrepant factor findings. Is it due to differences in sample characteristics (e.g., different ages)? Is it due to differences in the type of rater? Is it due to variability contributed by some problematic items included in the instrument? There are many possibilities. Fourth, suppose that a researcher retains a five-factor solution as most interpretable and meaningful, but one (or more) of the factors does not align well with the Lam and Aman model. This might be viewed as a partial replication by some, as the same number of factors were retained, and at least some (though not all) of the factors likely replicated. The individual factors that did replicate may come to be viewed as more robust, but the lack of full replication of the model will lead researchers to want to better understand why the variability in factor solutions is occurring across studies. Thus, after the initial EFA is conducted, it can be very instructive to assess to what extent prior factor solutions that are perceived to be important are at least partially present among the factor solutions extracted from a researcher's sample data.

Though ideas regarding the six- and five-factor solutions are implied, explicit hypotheses were not made for research questions four and five. The similarities and differences between the target factor solutions will be examined in terms of factor names/interpretations and item content.

Table 6. Summary of Study One Research Questions

Research Question	Research Question	Hypothesis	Analysis	Method(s) Used
1	When special education staff rate children with ASD, how many possible or likely interpretable factors are available for retention consideration?		EFA with PAF	*Scree plot, Kaiser criterion, Velicer's MAP, parallel analysis (PA)*

Table 6 (cont'd)

2	When special education staff rate children with ASD, how many factors should be retained to derive the most interpretable factor solution?	***2a.*** *The most interpretable EFA solution will consist of multiple factors* ***2b.*** *A SIB factor will be present.* ***2c.****, A Compulsive Behavior factor will be present*	EFA interpretative procedure	*Four researchers independently interpreted the potential solutions and retained the most interpretable factor solution by consensus.*
3	When special education staff rate children with ASD, does the most interpretable factor structure yield substantive correlations between at least some of the factors of the RBS-R?		EFA with oblique rotation	*Examination of the inter-factor correlation matrix for determination of the presence of substantive correlations*
4	When special education staff rate children with ASD, to what extend does the six-factor RBS-R solution from the present EFA (regardless of whether this was the retained solution) correspond to the six-subscale structure originally proposed by Bodfish et al. (2000)?		Qualitative comparison, calculation of the percentage of overlapping items per factor	*Examination of the factor constructs of the six-factor solution, compared to the Bodfish et al. (2000) six-factor solution*
5	When special education staff rate children with ASD, to what extent does the five-factor RBS-R solution from the present EFA (regardless of whether it was the retained solution) correspond to the five-factor solution reported by Lam and Aman (2007)?		Qualitative comparison, calculation of the percentage of overlapping items per factor	*Examination of the factor constructs of the five-factor solution, compared to the Lam and Aman (2007) five--factor solution*

Study One Sample Demographics

Participants included in the sample for study one consisted of 234 individuals with a diagnosis of ASD. Sample participants included 78.21% males (*n* = 183), 20.51% females (*n* = 48), and 1.28% unknown/missing (n = 3). Participants ranged in age from three to 21 years (*M* = 9.03, *SD* = 5.09; See Table 7). Racial and ethnic identification for this sample includes 73.50% White/non-Hispanic (*n* = 172), 11.54% Black (*n* = 27), 7.26% Hispanic/Latinx (*n* = 17), 2.99% Asian American, 0.43% Native American/Pacific Islander (*n* = 1), 1.71% "Other," (*n* = 4), and 2.56%% unknown (*n* = 6). Socioeconomic data were not available at the individual participant level for the study. However, agency-level data from the dates included in this study indicate that 29 - 36% of students qualified for a free or reduced lunch each year.

Cognitive deviation quotient (DQ) scores ranged from 15 to 127 (*M* = 59.98, *SD* = 21.57), with 69.9% of the sample presented with DQ scores $\leq$ 70 and 9.1% of the sample with DQ scores $\leq$ 85. DQ scores greater than 70 and 85 represent at least two and one standard deviations below the mean, respectively. All individuals included in the study sample presented with some level of severe functional impairments in domains of cognitive, social, and/or communication skills, as indicated by their placement in the agency's special education classrooms.

Table 7. Demographic Characteristics of Study One Sample

	Sample *N* (%)	Mean (*SD*)	Range
Participant Gender			
Male	183 (78.21%)		
Female	48 (20.51%)		
Missing	3 (1.28%)		
Participant Race/Ethnicity			
White/Non-Hispanic	172 (73.50%)		
Black	27 (11.54%)		
Hispanic/Latinx, no race specified	17 (7.26%)		

Table 7 (cont'd)

Asian American	7 (2.99%)		
Native American/Pacific Islander	1 (0.43%)		
Other	4 (1.71%)		
Unknown	6 (2.56%)		
Participant Age (years)		9.029 (*5.088*)	3.16–21.77 (18.61)
Participant Deviation Quotient Score		59.981 (*21.567*)	15.00-127.00 (112.00)
Unknown			

Note: All cognitive scores were set to a deviation quotient (DQ) metric (i.e., normative mean of 100, standard deviation of 15) to allow for limited comparability of cognitive scores

Data Analysis for Study One

All analyses for study one were conducted using several statistical programs. Programs included SAS Version 9.4 (SAS Institute Inc., 2013) and SPSS Version 26 (IBM Corp., 2019). Additionally, an R programming plugin was used in tandem with SPSS (Basto & Pereira; R Core Team, 2013). The primary data management system used for RBS-R data input is SPSS Version 26. All descriptive statistics was calculated using SPSS Version 26. The SPSS plugin for R will be utilized for conducting several more specialized analyses, including the generation of the inter-item polychoric correlation matrix, conducting a parallel analysis (Horn, 1965) and performing Velicer's MAP test (Velicer, 1976) calculations, and estimating internal consistency via Cronbach's alpha and ordinal alpha coefficients. Finally, SAS version 9.4 was used to run the EFA on the RBS-R inter-item polychoric correlation matrix, as generated from the SPSS plugin for R.

Pre-Analysis Data Cleaning and Missing Data

For the first study, data cleaning procedures and methods reflect those outlined by Osborne and Banjanovic (2016). Although all RBS-R protocols were previously completed, collected, and checked for missing items, some missing item ratings may have been missed. The frequency of missing data was not high enough to warrant biased analyses concerning missing data.

Correlation Matrix Sufficiency for Factoring

The 43 primary RBS-R items include ratings using a 0 to 3 scaling metric, which reflects an ordinal scale. Given the ordinal nature of the item variables, a polychoric correlation matrix was used to describe inter-item relationships, rather than a Pearson correlation matrix. Polychoric correlations attempt to correct for the range restriction that occurs when theoretically continuous constructs are measured using discrete ordered categories (Osborne & Banjanovic, 2016). Test items with fewer scale points artificially limit variability such that the use of Pearson correlations is not ideal coefficients and could underestimate the strength of relationships between the rating variables, ultimately biasing factor loadings (Kline, 1994; Osborne & Banjanovic, 2016).

Bartlett's Test of Sphericity (Bartlett, 1950) will be used to test the hypothesis that the inter-item correlation matrix is an identity matrix (i.e., with all true correlations of 0 and any non-zero correlation values due to chance; Pedhazur & Schmelkin, 1991). This is the first step, a minimal method, for establishing the presence of significant correlations in the input matrix for the proposed factor analysis.

In addition to Bartlett's Test, the KMO (Kaiser, 1960; Kaiser & Rice, 1974) test was performed on the correlation matrix, which assesses the extent of common variance present in the correlation matrix. The higher the KMO value the greater the amount of common variance in the matrix from which to derive factors. Kaiser and Rice (1974) provide KMO benchmark values for indicating data matrix suitability for factor analysis. Generally, KMO values below .5 indicate that matrices are not acceptable for an EFA, while KMO values above .8 indicate very suitable data matrices for an EFA. In greater detail, KMO values $\geq$ .90 are considered as "marvelous," values ranging from .80 to .89 as "meritorious," values ranging from .70 to .79 as "middling," values ranging from .60 to .69 as "mediocre," values ranging from .50 to .59 as "miserable," and values < .50 as "unacceptable," (Kaiser & Rice, 1974).

Within the literature, an agreement upon a strict sample size requirement is not yet been reached, but the literature typically considers that the larger the sample, the better (e.g., n = 300), providing a certain level of flexibility for smaller samples when there is strong data that yields multiple high factor loadings (e.g., .80 or greater; Costello & Osborne, 2005; Yong & Pearce, 2013). Guidelines that are more traditional indicate a ratio of four or five subjects per variable and use a sample of at least 200 subjects (Floyd & Widman, 1995). Further, MacCallum et al. (1999) provide guidelines regarding the likelihood of discovering an accurate factor solution, based on a simulation study conducted by researchers. If communalities and indicators in the model and study (i.e., number of factors, items, and participants) are known, researchers can estimate whether they have a large enough sample size to support an accurate factor analysis solution. Given the recommendations provided by MacCallum et al. (1999), the ratio of variables to factors for the RBS-R (Bodfish et al., 2000) calls for a sample size between 100 to 200 participants to be theoretically sufficient to support a factor solution. The total participants in this sample size (n > 200) are of moderate size and item distributions were anticipated to be non-normal, using the ASD sample.

Extraction Methods

Based on methods reported in more recent existing studies conducting factor analyses of the RBS-R with individuals with ASD (Georgiades et al., 2010; Russell et al., 2019), it is very likely that the RBS-R item data, involving relatively infrequent behaviors and based on cases with ASD, will violate univariate and multivariate normality assumptions. Consequently, PAF extraction, being very robust to such conditions, is the method of choice for the EFA in this study. Under these conditions, maximum likelihood (ML) extraction, the preferred EFA method for continuous, normally distributed data, would be inappropriate due to its strong normality assumption (Floyd & Widaman, 1995; Obsborne & Banjanovic, 2016) required for unbiased estimation.

Number of Factors to Retain

Determination of the most appropriate number of factors to retain in study one combines the use of the Guttman-Kaiser criterion (Guttman, 1954; Kaiser, 1960), the scree test (Cattell, 1966), parallel analysis (Horn, 1965), and the MAP test (Velicer et al., 2000). Utilization of these methods will aid in the identification of a range of factor solutions to examine for interpretability. Ultimately, the factor solution that is most interpretable and meaningful, from among these options, was retained.

Decision rules for the projected number of factors to interpret and retain are described in what follows. For the first method used, the Guttman-Kaiser criterion (Guttman, 1954; Kaiser, 1960), the rule states that all factors with eigenvalues greater than one will be retained for interpretation. Second, the scree test (Cattell, 1966) suggests the number of factors present through an examination of the graph of eigenvalues. The researcher looks for the natural bend (i.e., “elbow”) in the eigenvalue plot where the slope of the curve prominently flattens. The number of factors corresponding to the “elbow” in the scree graph is the number of factors suggested for interpretation and retention. Third, in the context of parallel analysis (Horn, 1965), factors are considered for retention if their obtained eigenvalues are significantly above the 95th percentile of the distribution of the random eigenvalues calculated via random re-ordering of the data (Ledesma & Valero-Mora, 2007; Velicer et al., 2000). Finally, for the MAP test (Velicer et al., 2000), common variance corresponding to each factor is partialled out as each successive factor emerges until only unique variance remains. When the common variance is depleted, the MAP value will be at its lowest (Osborne & Banjanovic, 2016).

Rotation

In study one, an oblique rotation method was used, as it was expected that all factors would be correlated, based on previous EFAs (Bishop et al., 2013; Georgiades et al., 2010; Kästel et al., 2020; Lam & Aman, 2007; Russell et al., 2019) regarding the RBS-R. It is noteworthy that within the factor analytic literature, oblique rotations are considered often as effective rotation strategies for both correlated and uncorrelated factors (Costello & Osborne, 2005). Therefore, an oblique direct oblimin rotation was implemented.

Interpreting the Solution

For study one, all factor loadings > .30 were considered significant (Beavers et al., 2013). Beyond this standard, all factor loadings $\geq$ .32 were considered as poor, $\geq$ .45 considered fair, $\geq$ .55 as good, $\geq$ .63 as very good, and $\geq$ .71 as excellent (Comrey & Lee, 1992). Any items loading $\geq$ .30 on more than one factor (i.e., cross-loadings) were thoroughly scrutinized to conclude which factor best reflects the rudimentary concept (Costello & Osborne, 2005).

Following statistical interpretation, factors underwent the factor naming process to aid in overall measurement understanding. Factors are named typically through consideration of what their most salient manifest variables share (Watkins, 2018). More specifically, based on guidelines outlined by Pett et al. (2003), the highest loaded items should offer a strong indication of the core of the emerging factor, especially true for loadings $\geq$ .90. However, loadings $\leq$ .60 tend to cause less robust interpretations (Pett et al., 2003). For study one, factor naming procedures followed recommendations previously listed (Pett et al., 2003; Watkins, 2018), while considering relevant aspects of repetitive behaviors most relevant for individuals with ASD. As part of study one, four investigators with knowledge of repetitive behaviors, ASD, and anxiety disorders independently examined the range of factor solutions, named factors, determined solutions, came together to share their interpretations, assessed similarities in interpretations, and discussed consensus regarding the most interpretable and clinically meaningful factor solution for retention.

Internal Consistency

Internal consistency reliability estimates for study one were examined using the collected RBS-R data. In this study, both ordinal alpha and Cronbach's alpha were calculated and reported. Ordinal alpha was used as the primary estimate of internal consistency, as it attempts to adjust for the ordinality of the item scales by replacing the commonly used Pearson correlations with polychoric correlations in the formula for coefficient alpha. Thus, the ordinal alpha estimates better fit the nature of the measurement scale involved (Gadermann et al., 2012). Estimates of Cronbach's alpha were also reported for comparison purposes—as the difference between the two types of alpha estimates makes clear the impact of the ordinal adjustment and allowed one to compare alpha results from the present study with results from prior studies that may have reported only a standard Cronbach's alpha.

In terms of interpretive benchmarks, Nunnally (1978) recommended the now widely used standards of a minimum internal consistency of .70 for research purposes and .90 or higher for important decisions about individuals. Salvia et al. (2017) suggested that measures used for high-stakes situations (e.g., special education decision-making) should have reliability coefficients of

$\geq$ .90, measures used for screening purposes should have a reliability of $\geq$ .80, and measures used for frequent progress monitoring should have reliability coefficients of $\geq$ .70.

Comparing Factor Solutions

Following the EFA, the six-factor solution extracted and rotated as part of the EFA and was compared to the original six-subscale RBS-R interpretive model proposed by Bodfish et al. (2000). The five-factor solution extracted and rotated during the EFA was also compared to the five-factor model reported by Lam and Aman (2007). Within this process, first, the factor/ subscale labels and interpretations were qualitatively compared across each pair of models for similarity. Following the qualitative comparison, the highest loading items were then compared across established and new factor solutions. Finally, the percentage of overlapping items between the previously published and newly obtained EFA-generated factor solutions were then assessed for agreement.

Study Two: CFA

Research Questions, Rationales, and Hypotheses

Research Question 6. When special education staff rate students with ASD, does the model generated by the EFA of the RBS-R in the exploratory sample fit the inter-item covariance matrix of the CFA validation sample?

Research Question 7. In the confirmatory sample, where special education staff rate students with ASD, how does the factor solution generated by the EFA of the RBS-R compare in terms of absolute and relative fit to the previous RBS-R factor models found in ASD samples (e.g., Hooker at al., 2019; Martinez-Gonzalez & Piqueras, 2017; Mirenda et al., 2010)?

Research Rationales and Hypotheses 6 and 7. Existing studies examining the factor structure of the RBS-R using CFA have yielded mixed results regarding the number of factors most supported for the instrument. Mirenda et al. (2010) examined six competing structural models existing in the literature at the time of the study (APA, 2000; Bodfish et al., 2000; Lam & Aman, 2007; Lam et al., 2008; Szatmari et al., 2006), regarding repetitive behavior. Results of the study suggested that both the five- and six-factor models showed evidence of a better statistical fit than the three- and four-factor models. The five- and six-factor models were almost identical regarding fit statistics, but the five-factor solution was selected instead, as it was more parsimonious than the six-factor (Mirenda et al., 2010). Hooker et al. (2019) examined six different factor models, yielding results that provided the most support for the four-, five-, and six-factor models. The four-factor model was not chosen, as it resulted in a factor correlation of .88 between two factors, indicating significant overlap between the two factors (Hooker et al., 2019), leaving the five- and six-factor models. For this study, the five-factor model demonstrated the most parsimonious, as the five- and six-factor models resulted in similar overall model fit indices. Of note, He et al. (2019) considered models consisting of between three and six factors using the Chinese RBS-R and retained a five-factor solution. He et al. (2019) found support for both a five- and six-factor model based on statistical fit but ultimately settled on the five-factor solution based on fit indices and model parsimony.

The present study's factor solution retained through an EFA in study one was hypothesized to be the most robust when compared to existing factor models for the RBS-R, because of thoroughness (i.e., using the most effective factor solution criterion methods, analyzing a range of potential factor solutions) of the analyses performed. Additionally, the exploratory sample and confirmatory sample from studies one and two are largely the same, in terms of ASD cases and rater types. Further, past studies' CFA models (i.e., Hooker et al., 2019; Mirenda et al., 2010) included different ASD samples (i.e., different ages), did not include raters beyond caregivers (i.e., school staff or teacher raters), and included models from different cultures and languages. Given this information, the model resulting from the present studies was likely to fit based on the consistent sample type, rater type, and language base of the RBS-R.

It was hypothesized that the RBS-R factor model selected as most interpretable in the EFA conducted in study one would at least adequately fit the RBS-R variance-covariance matrix of the second ASD sample when appropriately constrained for CFA. In this process, parameters for theoretically non-loading items were fixed to zero and a diagonally-weight least squares estimation procedure (i.e., the Weighted Least Squares Mean-Variance [WLSMV] estimator available through Mplus) was used given the ordinal nature of the data. To test the fit of the model, a combination of absolute, complexity-adjusted, and relative fit indices (e.g., adjusted chi-square, RMSEA, CFI, TLI, and Standard Root Mean Square Residual [SRMSR]) were used.

It was also theorized that the RBS-R factor model retained in the EFA conducted in study one would demonstrate a better fit to the second ASD sample's RBS-R inter-item variance-covariance matrix than the previously reported RBS-R factor models from the research literature when appropriately constrained for CFA. Researchers planned that the Mplus DIFFTEST (a variation on an adjusted chi-square test) would be used to test the significance of the difference in fit between two nested models if nesting assumptions were met. However, in the case of comparing non-nested models, the AIC and BIC fit indices were used for cross-model comparisons. Though the DIFFTEST is available through the WLSMV estimator in Mplus, the AIC and BIC indices, unfortunately, are not available using this estimation procedure. Because of this, the Mplus Robust Maximum Likelihood (MLR) estimator was implemented, but only to produce AIC and BIC values for cross-model comparisons.

Table 8. Summary of Study Two Research Questions

Research Question	Research Question	Hypothesis	Analysis	Methods Used
6	When special education staff rate children with ASD, does the mode; generated by the EFA of the RBS-R in the exploratory sample fit the inter-item covariance matrix of the CFA validation sample?? ***6.** The study one EFA solution will reasonably fit the confirmatory sample inter-item covariance matrix.*	*The study one EFA solution will reasonably fit the confirmatory sample inter-item covariance matrix.*	CFA with WLSMV estimator	*$\chi 2$, SRMR, RMSEA, CFI, TLI*
7	In the confirmatory sample, where special education staff rate children with ASD, how does the factor solution generated by the EFA of the RBS-R compare in terms of absolute and relative fit to the previous RBS-R factor models found in ASD samples (e.g., Hooker at al., 2019; Martinez-Gonzalez & Piqueras, 2017; Mirenda et al., 2010)? ***7.** The study one EFA model will demonstrate a better relative fit to the inter-item covariance matrix of the CFA sample when compared to previously reported RBS-R factor models derived from ASD samples.*	*The study one EFA model will demonstrate a better relative fit to the inter-item covariance matrix of the CFA sample when compared to previously reported RBS-R factor models derived from ASD samples. .*	CFA with WLSMV and MLR estimator	Mplus DIFFTEST for comparing nested models (if appropriate) and AIC and BIC indices for comparison of non-nested models

Study Two Sample Demographics

The sample for study two included 233 individuals with ASD. Participant demographics include 77.25% males (*n* = 180), 21.89% females (*n* = 51), and two missing/unknown (0.86%). Participants ranged from ages three to 22 years (*M* = 8.87, *SD* = 4.91; See Table 9). Individuals included in the study were identified as 78.11% White (n = 182), 8.58 % Black (n = 20), 6.01% Hispanic/Latinx (n = 14), 2.15% Asian American (n = 5), 2.58 % other (n = 6), 2.58 % unknown (n = 6). All socioeconomic data reported in study one is consistent with data for study two. Cognitive DQ scores range from 16.00 to 120.00 (M = 60.524, SD = 4.913), with 66.20 % of sample DQ scores $\leq$ 70 and 9.79% of sample DQ scores $\leq$ 85 (i.e., at least one standard deviation below and above the mean, respectively). Like study one, all participants included in study two present with substantial functional impairments in the cognitive, social, or communication domain, warranting participation in special education classrooms.

Table 9. Demographic Characteristics of Study Two Sample (N = 233)

Participant Demographics	Sample *N* (%)	Mean (*SD*)	Range
Gender			
Male	180 (77.25%)		
Female	51 (21.89%)		
Unknown	2 (0.86%)		
Race/Ethnicity			
White/Non-Hispanic	182 (78.11%)		
Black	20 (8.58%)		
Hispanic/Latinx, no race specified	12 (6.01%)		
Asian American/Pacific Islander	5 (2.15%)		
Other	6 (2.58%)		
Unknown	6 (2.58%)		
Age		8.871 (4.913)	3.06-22.31 (19.26)
Unknown	2 (0.86%)		
Deviation Quotient Score		60.524 (21.301)	16–120 (104)
Unknown	8 (3.43%)		

Note. All cognitive scores were set to a deviation quotient (DQ) metric (i.e., normative mean of 100, standard deviation of 15) to allow for limited comparability of participants' cognitive scores.

Data Analysis for Study Two

All analyses for study two were conducted using SPSS Version 26 (IBM Corp, 2019) and Mplus Version 8.6 (Muthén & Muthén, 1998-2017). SPSS Version 26 was primarily used for data management, which included inputting item-level RBS-R data and generating descriptive statistics for sample demographics and RBS-R item distributions. Mplus Version 8.6 was used for assessing the factorial validity of first-order confirmatory factor analytic models for the RBS-R. Primarily, the model of interest for study two stems from the study one EFA results. However, the model was compared to several other models previously reported in the literature that were generated in samples consisting of participants with ASD. The primary parameter estimates, and fit indices were generated using the Mplus WLSMV estimator. When comparing nested factor models, the Mplus DIFFTEST, available through the WLSMV estimation procedure, was used to assess the difference for significance. For purposes of comparing the relative fit of non-nested models, information criteria indices (e.g., AIC and BIC) were generated for each model using the Mplus robust maximum likelihood (MLR) estimator.

Pre-Analysis: Data Cleaning and Missing Data

Data cleaning processes in study two followed the same procedures outlined in study one. Again, all RBS-R protocols were previously completed, collected, and checked for missing items, some missing item ratings may have been missed. The frequency of missing data was not expected to be high enough to warrant biased analyses concerning missing data.

Correlation Matrix Sufficiency for Factoring

As with EFA, the CFA literature involves some disagreement and ongoing investigation of the minimum and ideal sample sizes required for a CFA (MacCallum et al., 1999; Marsh et al., 2009; Schmitt, 2011). Most articles emphasize the need for larger sample sizes to best support the conduction of a CFA and yield stable results (Harrington, 2009; MacCallum et al., 1999; Marsh et al., 2009; Schmitt, 2011). MacCallum et al. suggest that the same ratio of variables to factors with moderate to high communality estimates acceptable for EFA should be also acceptable for CFA. Therefore, a sample size between 100 and 200 would theoretically be sufficient for supporting convergent solutions in study two, based on calculations from study one. Additional research has suggested that non-normally distributed data calls for increased sample size, changing from 150 for a normal distribution to 265 required for the non-normal distribution of data (Muthén & Muthén, 2002). Total participants in this sample size ($n > 200$) are of moderate size and item distributions are non-normal, using the ASD sample, though not perfect by these standards.

The existing dataset for study two was examined for selecting the most appropriate estimation method for conducting the CFA, regarding multivariate normality (i.e., multivariate normal or multivariate non-normal). Given the ordinal nature of the RBS-R data (i.e., four-point Likert-type scale items are ordinal in nature) and the use of measurement data from an ASD sample (e.g., similar to Birnbaum, 2020 and Mirwis, 2011), data were expected to be non-normal. Therefore, it was anticipated that a robust diagonally weighted procedure was most fitting for this study. The use of the WLSMV estimator addresses this issue, though it is a concern that extreme non-normality in the data can risk altering standard errors and statistical power (DiStefano & Morgan, 2014). However, in previous studies with extremely non-normal data, average RMSEA and CFI values did not appear to be sensitive to such differences in normality, and simulations have shown the Mplus WLSMV procedure is preferable to the LISREL diagonally weighed estimation option for moderate non-normal data with smaller sample sizes (DiStefano & Morgan, 2014).

Model Specification

Model specification techniques for CFA involve detailing the specific models to be tested, including specification of the observed and latent variables, unique variance (i.e., error variance in each item not accounted for by the latent factors), inter-factor correlations, and directional paths from factors to items (Harrington, 2009). To best illustrate the model specification process, a graphical structure is used to denote the paths and parameters, depicting relationships among the variables. The items, or observed variables, are represented by rectangles, while the factors, or latent variables, are represented by ovals. Single-headed arrows represent all directional paths between latent and observed variables, while double-headed arrows (Byrne, 2012; Harrington, 2009) represent correlations between the variables. Within CFA graphs, arrows going from latent to observed variables represent the effect of the latent variables or constructs on observed variables. Within the graph, factor loadings are provided for each variable, comparable to regression coefficients, which predict the observed variables from the unobserved factors (Harrington, 2009). Error terms (i.e., referred to as residuals in Mplus) are represented by a direct path arrow, starting from the identified error term, and pointing to an observed variable, reflecting measurement error (i.e., random error and unique variance not accounted for by the latent variables). Identified error terms typically have set paths fixed to 1.0 to provide a scale for the error term (based on the items) and have variances freely estimated (Byrne, 2012).

Several models were analyzed for the CFA in study two. The primary model of interest was the model derived from the EFA in study one, which was assessed alongside models derived from previous factor analyses of the RBS-R reported in the research literature. Additional models included the two-factor model (Georgiades et al., 2010 [Greek RBS-R]), four-factor model

(Russell et al., 2019; Kästel et al., 2020 [German RBS-R]), five-factor models (Bishop et al., 2013; Lam & Aman, 2007; Sturm et al., 2022), and six-factor model (Bodfish et al., 2000; Martínez-González & Piqueras, 2017 [Spanish RBS-R]). Prior evidence for all existing models was derived from samples of participants with ASD, rated by parents or caregivers, except for one six-factor model (Martínez-González & Piqueras, 2017), which included school staff ratings. It is noteworthy that the samples for studies one and two in the present project include ratings from special education staff and teachers only and not from parents.

Model Identification

The process of model identification in a CFA includes providing a scale for each latent variable in the model (i.e., the establishment of a measurement unit for the latent variables) and guaranteeing the degrees of freedom (*df*) in the model are > 0 (Harrington, 2009). To estimate model parameters, there must be more unique information elements in the variance-covariance matrix than there are unknown parameters to be estimated in the factor model. In the case of more unknown parameters to be estimated than there are elements in the variance-covariance matrix, the model cannot be properly estimated due to insufficient *df*. Primarily, the *df* signifies differences between the total information elements available in the inter-item variance-covariance matrix and the unknown parameters to be freely estimated. Depending on the value of the *df*, models can be under-identified (i.e., $df < 0$), just identified ($df = 0$), or over-identified ($df > 0$; Harrington, 2009). For study two, all models were over-identified.

To assign meaningful units of measurement for latent variables, scaling the items is an essential task to perform in a CFA (Harrington, 2009). In study two, the fixed factor method was utilized to allow all factor loadings to be freely estimated. The fixed factor method consists of setting all factor variances to 1.0, allowing all factor loadings to be freely estimated using factor variance units (Byrne, 2012). By using the fixed factor method, the interpretability of the inter-factor covariances improves, as they can be interpreted as correlation coefficients following their standardization.

Model Estimation

To establish whether the hypothesized model is congruent with the variance-covariance data, model parameters (e.g., factor loadings, error variances, inter-factor covariances) must be estimated to determine the quality of the fit. In doing so, calculations are repeated, increasing precision with each iteration until the convergence criterion is reached and the model is as precise as possible (Harrington, 2009). For this study, the WLSMV (Muthén & Muthén, 1993; Muthén et al., 1997; Muthén & Muthén, 2017) estimator was used with the polychoric correlation matrix and sample estimated asymptomatic covariance matrix as input, given that the item data are ordinal. These methods are similar to those used by Hooker et al. (2019) for conducting a CFA on the RBS-R in a young ASD sample rated by parents/caregivers.

Model Fit

Following the completion of model estimation methods on the hypothesized model(s), model fit methods were utilized for assessing how well models align with the data. At present, there are limited recommendations regarding a standardized process to follow for assessing model fit (Brown, 2006; Iacobucci, 2010; Jackson et al., 2009). However, it is recommended that one use at least one fit index each across three fit index categories, including absolute fit indices, fit adjusted for model parsimony, and comparative (i.e., incremental) fit indices (Brown, 2006). Furthermore, researchers (Jackson et al., 2009) suggest using a chi-square value with *df* and probability value, an incremental fit index (i.e., comparative fit index), and residuals-based measures within the model fit process.

For study two, the WLSMV-adjusted chi-square absolute fit index and the SRMR were used for examining the relationship between the predicted and sample variance-covariance matrices. The chi-square indicates if the model of interest sufficiently replicates variances and covariances present in the sample data (Brown, 2015). In the instance of a statistically significant chi-square value, the model does not represent an acceptable fit with the sample data. However, the SRMR method was used concurrently, as the chi-square method is more vulnerable to sample size and non-normal distribution of data—where non-substantive deviations can be declared statistically significant by an overpowered hi-square (Brown, 2015). The SRMR was used to measure the differences between the data matrix and correlations emerging from the hypothesized model (Harrington, 2009), represented by how discrepant the model is from a perfect fit (i.e., the value of 0). Though guidelines for model fit are limited, Hu and Bentler (1999) suggest a recommended cutoff value of .08 for the SRMR, while values can range from zero to one. For this index, lower SRMR values are preferred, indicating a better fit (Hu & Bentler, 1999).

As an additional measure of model fit, study two includes the use of parsimony collection indices, which take into consideration the number of *df* (Brown, 2006). To do so, this study will include RMSEA (Steiger, 2016; Steiger & Lind, 1980), the AIC (Akaike, 1987), and the BIC (Kass & Raftery, 1995) as parsimony correction indices. The RMSEA is included in study two for its lack of sensitivity to sample size as it estimates the degree of model fit relative to the population (Brown, 2006). Like with the SRMR, a perfect fit is indicated by a value of zero, and the fit is assessed by how close the calculated value is to 0. Browne and Cudeck (1992) list guidelines for interpretation, including values $\leq$ .05 as a “close fit,” values between .05 and (equal to) .08 as a “reasonable fit,” and values $\geq$ .10 as not a good fit for the model. Additionally, Hu and Bentler (1999) proposed a cutoff of .06 for model fit. The AIC and BIC indices for parsimony correction were also included in study two to aid in the comparison of the non-nested models being assessed within the same dataset. The AIC and BIC assign penalties to model fit based on model complexity. Generally, the lower the value of the AIC and BIC the more satisfactory the hypothesized model (Byrne, 2012; Harrington, 2009), though there are no specific cut-scores for these. They are most useful for comparing models.

Finally, comparative fit indices are used frequently to assess the fit of a hypothesized model, when compared to a more restricted, nested model. In this case, the restricted model leaves out any covariance between observed variables, allowing the variables to remain independent (Brown, 2006). Comparative fit indices allow for comparison between the hypothesized model and a simpler version of itself, barring any correlations between variables (Iacobucci, 2010). Study two will include the CFI (Bentler, 1990) and the TLI (Tucker & Lewis, 1973) for these purposes. Similar to previously discussed methods, CFI values range from 0 to 1 and values greater than .95 are considered a good fit, though the range from .90 to .95 is deemed fair for interpretation and consideration (Brown, 2006; Hu & Bentler, 1999). The TLI values are non-normed. Therefore, values may range from zero to above one, with values closer to one preferable for an acceptable model fit (Brown, 2006; Byrne, 2012). The TLI method includes a penalty for more complex models.

Direct comparison of the RBS-R author-proposed model (Bodfish et al., 2000) and the EFA-generated model from study one included the utilization of the AIC and BIC indices (Brown, 2015; Muthén & Muthén, 1998-2017). The AIC and BIC indices were used descriptively for cross-model comparisons. However, the WLSMV estimator does not allow for AIC and BIC calculations. Therefore, the MLR estimator was used to supplement the WLSMV to attain AIC and BIC values for the models (see Birnbaum, 2020). It was considered that in the case that one of the models is nested within another, the assumption would be met for use of the Mplus DIFFTEST (i.e., corrected chi-Square difference test available when using WLSMV). In the case that the models do not meet nesting assumptions, the DIFFTEST cannot be used.

CHAPTER FOUR: RESULTS

Study one included the analysis of the factor structure of the Repetitive Behavior Scale-Revised (RBS-R; Bodfish et al., 2000). The sample for this study included individuals with a diagnosis of autism spectrum disorder (ASD) and utilized ratings by special education teaching staff familiar with the participants, making it unique from other studies using the RBS-R that had only used parent ratings. The purpose of study one was to determine the number of potentially interpretable factors for retention and to examine the presence of substantive correlations between the factors. Additionally, the first study sought to assess the overlap – in terms of factors and factor item content that existed between the factor model from the study one exploratory factor analysis (EFA) and the six-subscale (i.e., six-factor) model proposed by the test authors (Bodfish et al., 2000). Study one utilized an inter-item polychoric correlation matrix as input for the EFA with principal axis factoring (PAF) for extraction and a direct oblimin rotation to assess for substantive inter-factor correlations. Internal consistency reliability estimates were obtained using ordinal alpha as the primary estimation approach, and Cronbach's alpha, to provide supplemental estimates for standard for comparison with other studies that reported only Cronbach's alpha estimates. Study two examined the fit of both the model from study one and the published six-factor model, based on the RBS-R author's six-subscale structure, in a second sample of students with ASD (independent of the EFA sample from study one) with RBS-R ratings by special education teaching staff familiar with the participants who were rated. This study focused on absolute fit, fit adjusted for model parsimony, and comparative fit of the factor structures listed using confirmatory factor analysis (CFA). The CFA was run using the robust weighted least squares mean variance (WLSMV) estimation procedure available in Mplus considering the ordinal nature of the item scaling input. Results are presented for both studies in relation to study-specific subsets of the seven research questions that guided this dissertation.

Data Cleaning and Missing Data

The datasets for both studies one and two were scanned for missing values before performing the respective data analyses. Frequency statistics were run to assess for values falling outside the appropriate scale range and missing values; discrepant values were identified, corrected, and verified by checking the physical copy of the case record form. Missing values were identified and checked against the physical case record form. If a missing value was available on the physical case record it was entered into the database. Data were then assessed for the percentage of true missing values. The data set was assessed as missing values for 0.11% of items across all cases, equal to 22 total missing item values out of 20,081 item scores (i.e., 467 cases [i.e., total EFA and CFA cases combined] x 43 RBS-R items per case). For the nine cases identified as having missing values, the number of missing item values per case was between one and three missing values, with the mode being three missing items per case (M = 2.44). The expectation maximization (EM; Moon, 1996) method was used to predict and impute values for the missing items, allowing for their inclusion in the analyses for the two studies.

Study One: EFA

Data Matrix Sufficiency for Factoring

Descriptive statistics, including the mean, standard deviation, and scale response percentages, for each of the 43 RBS-R items for study one were evaluated and are presented in Table 10. Each item was rated on a four-point scale ranging from 0 “behavior does not occur” to 3 “behavior occurs and is a severe problem” (Bodfish et al., 2000). Item stems reported in all tables of this dissertation are truncated. It is noteworthy that the response distributions for all items are non-normal and similar (i.e., largest percentage of cases responding at the lowest scale value). This was expected given the extreme behaviors represented and their anticipated relatively low frequency even in an ASD sample. The inter-item polychoric correlation matrix is reported in Appendix A, which includes estimates of how each item relates to the others in the dataset. This matrix was used as input for the EFA.

Table 10. Study One Dataset Item-Level Descriptive Statistics (N = 234)

Item Content				Percent of Responses Per Item			
Item		Mean	Standard Deviation	0 Behavior does not occur	1 Behavior occurs and is a mild problem	2 Behavior occurs and is a moderate problem	3 Behavior occurs and is a serious problem
1	WHOLE BODY (Body rocking, body swaying)	0.812	0.997	51.3%	25.6%	13.7%	9.4%
2	HEAD (rolls head, nods head, turns head)	0.491	0.798	67.1%	19.7%	10.3%	3.0%
3	HAND/FINGER (Flaps hands, wiggles or flicks fingers, claps hands, waves or shakes hand or arm)	1.073	1.084	41.9%	22.2%	22.6%	13.2%
4	LOCOMOTION (turns in circles, whirls, jumps, bounces)	0.897	1.063	50.9%	19.7%	18.4%	11.1%
5	OBJECT USAGE (spins or twirls objects, twiddles or slaps or throws objects, lets objects fall out of hands)	0.927	1.056	48.3%	21.8%	18.8%	11.1%
6	SENSORY (covers eyes, looks closely or gazes at hands or objects, covers ears,	1.013	1.070	44.0%	23.1%	20.5%	12.4%

Table 10 (cont'd)

	smells, or sniffs items, rubs surfaces)						
7	HITS SELF WITH BODY PART (Hits or slaps head, face, or other body area)	0.633	1.024	66.7%	14.1%	8.5%	10.7%
8	HITS SELF AGAINST SURFACE OR OBJECT (hits or bangs head or other body part on table, floor, or other surface)	0.577	0.987	69.2%	13.2%	8.1%	9.4%
9	HITS SELF WITH OBJECT (Hits or bangs head or other body area with objects)	0.333	0.781	81.2%	9.0%	5.1%	4.7%
10	BITES SELF (Bites hand, wrist, arm, lips, or tongue	0.419	0.841	76.5%	9.8%	9.0%	4.7%
11	PULLS (pulls hair or skin)	0.278	0.702	83.3%	9.0%	4.3%	3.4%
12	RUBS OR SCRATCHES SELF (Rubs or scratches marks on arms, leg, face, or torso)	0.368	0.719	75.2%	15.0%	7.7%	2.1%
13	INSERTS FINGER OR OBJECT (eye-poking, ear-poking)	0.145	0.495	90.2%	6.4%	2.1%	1.3%
14	SKIN PICKING (picks at skin on face, hands, arms, legs, or torso)	0.235	0.668	86.3%	7.3%	3.0%	3.4%
15	ARRANGING/ ORDERING (arranges certain objects in a particular pattern or place;	0.547	0.889	67.1%	16.2%	11.5%	5.1%

Table 10 (cont'd)

	Need for things to be even or symmetrical)						
16	COMPLETENESS (Must have doors opened or closed; Takes all items out of a container or area)	0.526	0.880	68.4%	15.8%	10.7%	5.1%
17	WASHING /CLEANING (Excessively cleans certain body parts; Picks at lint or loose threads)	0.239	0.644	85.0%	8.5%	3.8%	2.6%
18	CHECKING (Repeatedly checks doors, windows, drawers, appliances, clocks, locks, etc.)	0.201	0.606	88.0%	6.0%	3.8%	2.1%
19	COUNTING (Counts items or objects; Counts to a certain number or in a certain way)	0.141	0.542	91.9%	4.7%	0.9%	2.6%
20	HOARDING/SAVING (Collects, hoards, or hides specific items)	0.316	0.689	79.1%	12.4%	6.4%	2.1%
21	REPEATING (Need to repeat routine events; In/outdoor, up/down from chair, clothing on/off)	0.342	0.771	80.3%	9.0%	6.8%	3.8%
22	TOUCH/TAP (Need to touch, tap, or rub items, surfaces, or people)	0.483	0.875	72.2%	12.4%	10.3%	5.1%
23	EATING/ MEALTIME (Strongly prefers/insists on	0.944	1.146	51.3%	20.1%	11.5%	17.1%

Table 10 (cont'd)

	eating/drinking only certain things; Eats or drinks items in a set order; Insists that meal related items are arranged in a certain way)						
24	SLEEPING/ BEDTIME (Insists on certain pre-bedtime routines; Arranges items in room "just so" prior to bedtime; Insists that certain items be present with him/her during sleep; Insists that another person be present prior to or during sleep_	0.410	0.820	75.6%	12.4%	7.3%	4.7%
25	SELF-CARE - BATHROOM AND DRESSING (Insists on specific order of activities or tasks related to using the bathroom, to watching, showering, bathing, or dressing; Arranges items in a certain way in the bathroom or insists that bathroom items not be moved; Insists on wearing certain clothing items)	0.355	0.728	76.9%	13.2%	7.3%	2.6%
26	TRAVEL/TRANS-PORTATION (Insists on taking certain routes/paths; Must sit in specific location in vehicles; Insists that certain items be present during travel,	0.299	0.665	79.5%	13.2%	5.1%	2.1%

Table 10 (cont'd)

	e.g., toy or material; Insists on seeing or touching certain things or places during travel such as a sign or store)						
27	PLAY/LEISURE (Insists on certain play activities; Follows a rigid routine during play/leisure; Insists that certain items be present/available during play/leisure; Insists that other persons do certain things during play	0.620	0.882	60.3%	22.2%	12.8%	4.7%
28	COMMUNICATION /SOCIAL INTERACTIONS (Repeats same topic(s) during social interactions; Repetitive questioning; Insists on certain topics of conversation; Insists that others say certain things or respond in certain ways during interactions)	0.415	0.788	73.9%	14.1%	8.5%	3.4%
29	Insists that things remain in the same place(s) (e.g., toys, supplies, furniture, pictures, etc.)	0.470	0.819	70.1%	16.7%	9.4%	3.8%
30	Objects to visiting new places	0.423	0.762	71.4%	17.9%	7.7%	3.0%
31	Becomes upset if interrupted in what he/she is doing	0.992	0.945	37.2%	34.2%	20.9%	7.7%

Table 10 (cont'd)

32	Insists on walking in a particular pattern (e.g., straight line)	0.252	0.622	83.3%	9.4%	6.0%	1.3%
33	Insists on sitting at the same place	0.312	0.649	77.8%	14.5%	6.4%	1.3%
34	Dislikes changes in appearance or behavior of the people around him/her	0.235	0.608	84.2%	9.8%	4.3%	1.7%
35	Insists on using a particular door	0.120	0.386	89.7%	9.0%	0.9%	0.4%
36	Likes the same CD, tape, record, or piece of music played continually; Likes same move/video or part of movie/video	0.641	0.993	64.5%	15.8%	10.7%	9.0%
37	Resists changing activities; Difficulty with transitions	0.923	1.016	46.2%	24.8%	19.7%	9.4%
38	Insists on same routine, household, school, or work schedule everyday	0.564	0.907	66.2%	17.1%	10.7%	6.0%
39	Insists that specific things take place at specific times	0.342	0.749	79.1%	11.1%	6.4%	3.4%
40	Fascination, preoccupation with one subject or activity (e.g., trains, computers, weather, dinosaurs)	0.846	1.053	52.6%	21.8%	14.1%	11.5%
41	Strongly attached to one specific object	0.727	1.016	58.1%	21.8%	9.4%	10.7%
42	Preoccupation with part(s) of object rather than the whole object (e.g., buttons	0.509	0.875	69.2%	16.2%	9.0%	5.6%

Table 10 (cont'd)

	on clothes, wheels on cars)						
43	Fascination, preoccupation with movement/things that move (e.g., fans, clocks)	0.363	0.775	78.2%	11.1%	6.8%	3.8%

To address the suitability of the inter-item correlation matrix for an EFA (i.e., presence of significant correlations and common variance), Bartlett's Test of Sphericity (Barlett, 1950) and the Kaiser-Meyer-Oklin test of sampling adequacy (KMO; Kaiser, 1960; Kaiser & Rice, 1974) were examined. The Bartlett's test was significant, indicating that the correlation matrix was not an identity matrix (χ^2 = 5145.358, df = 903, p <0.001; Raykov & Marcoulides, 2008). The KMO test of sampling adequacy result was 0.861. Results aligned with Kaiser and Rice's (1974) KMO standards, which state that the KMO should, at minimum, be above 0.50. More specifically, Kaiser and Rice provided guidance concerning appropriateness for factor analysis of KMO values within various ranges (i.e., estimates falling between 0.50 and 0.59 are considered "miserable," 0.60 to 0.69 are "mediocre," 0.70 to 0.79 are "middling," 0.80 to 0.89 are "meritorious," and values above .90 are considered "marvelous" [Kaiser & Rice, 1974]). The KMO value .861 for study one met the second-highest standard (i.e., "meritorious") consistent with the presence of substantive common variance present for conducting factor analyses. Results of both Bartlett's Test of Sphericity and KMO test of sampling adequacy clearly established that the inter-item correlation matrix was sufficient to perform an EFA.

The adequacy of the sample size study one's EFA was evaluated using the simulation results under various conditions provided by MacCallum et al. (1999). RBS-R item communality estimates for study one's sample ranged from 0.454 to 0.769 and were considered wide (i.e., values generally ranging between .400 and .800). MacCallum et al.'s Monte Carlo simulation results were used to estimate the likelihood of a convergent factor solution based on the ratio of items to factors, communality estimates, and sample size. Based on the RBS-R estimated item-to-factor ratio (43:6), the closest corresponding table entry was a 20:3 ratio, which yielded 100% convergent admissible solutions in simulations at a sample size of 100 or above in the context of wide item communality estimates (see Table on p. 93 of MacCallum et al., 1999). Thus, according to the standards described, the 234-subject sample size for study one was suitable for the EFA.

Research Question 1

When special education staff rate students with ASD, how many possible or likely interpretable RBS-R factors are available for retention consideration?

Research question one was addressed using Principal Axis Factoring (PAF), the Guttman-Kaiser Criterion (Guttman, 1954; Kaiser, 1960), the scree test (Cattell, 1966), parallel analysis (Horn, 1965), and the Velicer's minimum average partial test (V-MAP: Velicer, 1976). PAF and the factor indices were estimated using a combination of SPSS, SAS, and R statistical packages. For the initial extraction, PAF was chosen based on the assumption that the dataset would likely violate univariate and multivariate normality—which it clearly did, based on item distributions noted in Table 10. The PAF was conducted in SAS using the inter-item polychoric correlation matrix provided by R package polycor (R3.1.2; Basto & Pereira, 2012; R Core Team, 2013).

Ultimately, the results of the four criteria for factor selection (i.e., Kaiser criterion, scree plot, parallel analysis, and Velicer's MAP) were used to provide a range of possible factor solutions (to answer research question one), which would then be examined for interpretability under research question two. As indicated in the method section (Chapter 3), the number of factors suggested by each criterion should be thought of as that value plus or minus one factor in terms of the number of factors to examine. Thus, for example, if parallel analysis suggested a four-factor solution for retention, this would suggest the need to examine the interpretability of factor solutions consisting of between three and five factors. This "plus or minus one" principle applied to the Kaiser criterion, scree plot, and Velicer's MAP, as well.

The Kaiser criterion (Kaiser, 1960) states that factors with eigenvalues greater than one should be retained. (It is noteworthy that this approach was based on principal components analysis, where "1.0" values are always placed in the diagonal of the inter-item correlation matrix. An equivalent metric in PAF would be something like the mean of all item communality estimates. However, the original eigenvalue greater than one rule is typically used in practice regardless of context. Both standards were applied here.) Based on the values obtained from the SPSS R plugin Ordinal Factor Analysis Menu (Basto & Pereira, 2012), results suggested that eight factors should be considered for retention. Table 11 lists all the observed eigenvalues generated from SPSS R Ordinal Factor Analysis menu and SAS outputs. Using the SAS values and eigenvalue greater than 1.0 criterion, six potential factors were suggested. To avoid overlooking any potential factor solutions, the interpretation of nine possible factor solutions was used given the exploratory nature of the analysis.

Table 11. Eigenvalues for the Guttman-Kaiser Criterion

Possible Factor	Eigenvalues (SPSS R Plugin)	Eigenvalues (SAS)
1	16.412	16.266
2	4.776	4.674
3	2.599	2.494
4	1.961	1.813
5	1.341	1.220
6	1.266	**1.096**

Table 11 (cont'd)

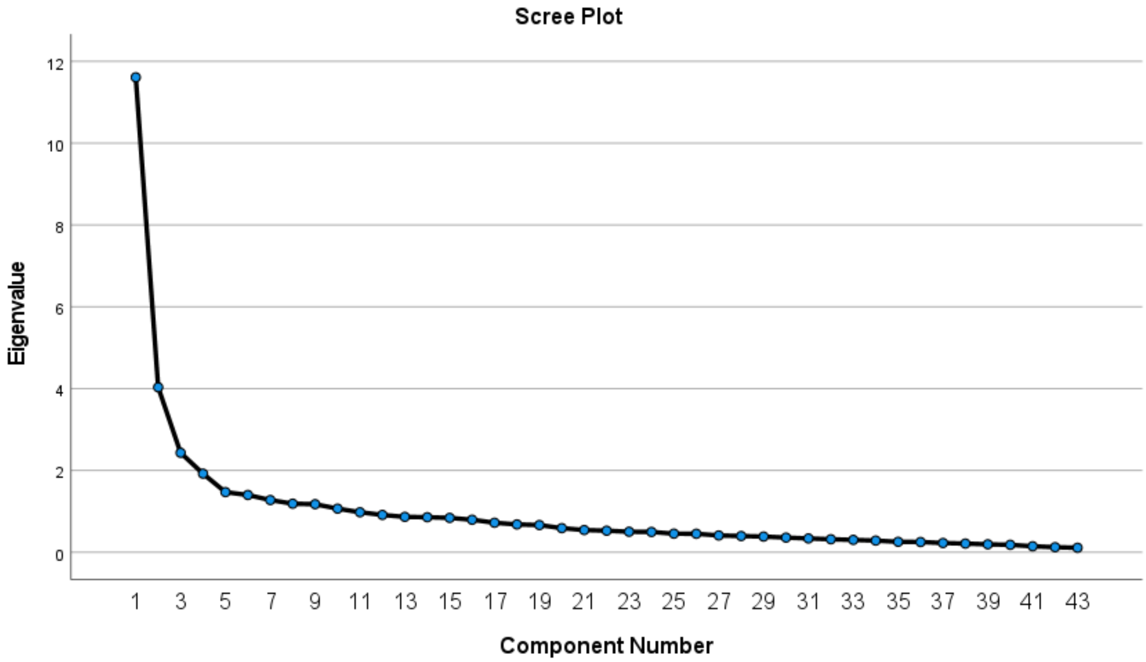
7	1.137	0.917
8	<u>**1.028**</u>[a]	0.877
9	0.970	0.749
10	0.882	0.594
11	0.777	0.561
12	0.707	0.521
13	0.688	0.487
14	0.610	0.425
15	0.581	0.372
16	0.487	0.305
17	0.408	0.263
18	0.377	0.229
19	0.324	0.175
20	0.302	0.142
21	0.257	0.092
22	0.200	0.038
23	0.177	0.014
24	0.138	-0.023
25	0.118	-0.043
26	0.095	-0.075
27	0.057	-0.088
28	0.029	-0.113
29	0.017	-0.118
30	0.010	-0.160
31	-0.017	-0.176
32	-0.030	-0.194
33	-0.047	-0.203
34	-0.071	-0.230
35	-0.085	-0.231
36	-0.094	-0.261
37	-0.133	-0.276
38	-0.207	-0.300
39	-0.234	-0.312
40	-0.254	-0.318
41	-0.312	-0.337
42	-0.386	-0.342
43	-0.446	-0.363

Note. The first eigenvalue above one is bolded and underlined in the table for each SPSS with R Plugin (Basto & Pereira, 2012; R Core Team, 2013) and SAS calculations.

[a] The SPSS Ordinal Factor analysis Menu access through the R Plugin calculates the eigenvalues differently from SAS. The Kaiser Criterion suggests 8 factors for the SPSS with R Plugin and 6 factors for the SAS calculation.

The scree test using eigenvalues generated from the SPSS R plugin can be found in Figure 1. The scree test shows a downward curving line with points that indicate eigenvalues. This test is interpreted by visually inspecting the slope of the line to determine when it becomes level. Once the identification of the "break" in the graph is determined, the next step is to look at factor solutions with the number of factors above the identified break. Visual inspection of the scree plot of the current study suggested a break or point of inflection of the line at the third and sixth eigenvalues. These results indicate that an examination of the factor solutions with the number of data points prior to these breaks – two- and seven-factor structures is warranted. Therefore, a two- and seven-factor solution were retained for consideration.

Figure 1. Scree Plot of Eigenvalues Generated by the SPSS R Plugin Ordinal Factor Analysis Menu



Parallel analysis (Horn, 1965) was calculated through the SPSS R Plugin as well. Parallel analysis is used to compare the variance of the extracted factors from the obtained dataset to variance of factors generated via random reconfigurations of the dataset (Horn, 1965; O'Connor, 2000). In this study, eigenvalues were generated based on 100 randomly generated samples resulting from the random re-arrangement of the data from the 234 cases. Observed eigenvalues were then compared to the randomly generated eigenvalues. Based on criteria set by Glorfeld (1995), observed factors with eigenvalues above the 95th percentile of the randomly generated eigenvalues were retained. The observed and randomly generated eigenvalues for all 43 potential factors are reported in Table 12. Results of the parallel analysis, using the 95th percentile criterion, suggested the retention of four factors. Observed eigenvalues beyond the first four fell below the values at the 95th percentile of the random eigenvalue distribution for that factor.

Table 12. Parallel Analysis with Observed and Random Eigenvalues at the 95th Percentile

Potential Factor	Observed Eigenvalues	Parallel Analysis Eigenvalues
1	16.412	2.409
2	4.776	2.238
3	2.599	2.025
4	**1.961**	**1.902**
5	1.341	1.816
6	1.266	1.722
7	1.137	1.598
8	1.028	1.495
9	0.910	1.441
10	0.882	1.356
11	0.777	1.245
12	0.707	1.170
13	0.688	1.114
14	0.610	1.037
15	0.581	0.992
16	0.487	0.942
17	0.408	0.912
18	0.377	0.872
19	0.324	0.762
20	0.302	0.717
21	0.257	0.660
22	0.200	0.645
23	0.177	0.584
24	0.138	0.537
25	0.118	0.467
26	0.095	0.403
27	0.057	0.337
28	0.029	0.262
29	0.017	0.230
30	0.010	0.189
31	-0.017	0.138
32	-0.030	0.079
33	-0.047	0.048
34	-0.071	0.030
35	-0.085	-0.004
36	-0.094	-0.075
37	-0.133	-0.079

Table 12 (cont'd)

38	-0.207	-0.153
39	-0.234	-0.198
40	-0.254	-0.257
41	-0.312	-0.313
42	-0.386	-0.395
43	-0.446	-0.807

Note. The parallel analysis eigenvalues reflect the 95th percentile of the eigenvalue distribution.

Velicer's MAP (Velicer, 1976) was generated from the SPSS R plugin Ordinal Factor Analysis Menu (Basto & Pereira, 2012). The MAP test includes an extraction of each consecutive factor and then provides an index of the remaining common variance. As the variance values begin to increase again following the factor with the minimum value, this indicates that the variance being extracted has moved to unique, rather than common variance (O'Connor, 2000). The suggested number of factors to retain is determined when the common variance of the factors reaches its minimum point, leaving only a unique variance (Osborne & Banjanovic, 2016). Table 13 displays the results from Velicer's MAP test for this study, including both average partial correlations and fourth average (i.e., a revision to the original MAP test analysis where partial correlations were raised to the fourth, rather than second, power to improve accuracy) partial correlations (Velicer et al., 2000). The results using the squared average partial correlations suggested the retention of four factors, as the fourth factor has the lowest squared average partial correlation of 0.0296. The results using the fourth average partial correlations suggested retention of six factors, as the sixth factor had the lowest fourth average partial correlation of 0.0027.

Table 13. Velicer's MAP Test Values for Squared Average and 4th Average Partial Correlations

Factors	Squared Average Partial Correlations	4th Average Partial Correlations
0	0.1562	0.0393
1	0.0491	0.0078

Table 13 (cont'd)

2	0.0352	0.0040
3	0.0314	0.0034
4	**0.0296**	0.0030
5	0.0300	0.0029
6	0.0308	**0.0027**
7	0.0325	0.0030
8	0.0335	0.0032
9	0.0351	0.0032
10	0.0380	0.0038
11	0.0405	0.0044
12	0.0434	0.0051
13	0.0462	0.0059
14	0.0496	0.0069
15	0.0526	0.0073
16	0.0552	0.0083
17	0.0592	0.0099
18	0.0641	0.0112
19	0.0676	0.0123
20	0.0702	0.0132
21	0.0755	0.0144
22	0.0822	0.0168
23	0.0920	0.0212
24	0.1018	0.0262
25	0.1142	0.0317
26	0.1250	0.0385
27	0.1466	0.0502
28	0.1749	0.0693
29	0.2061	0.0957
30	0.2642	0.1394
31	0.3769	0.2399
32	0.4411	0.3038
33	0.6902	0.5821
34	0.9209	0.8740
35	0.1111	0.0310
36	0.1326	0.0444
37	0.1593	0.0598
38	0.1968	0.0860
39	0.2465	0.1238
40	0.3250	0.1895
41	0.4944	0.3668

Clearly, the number of suggested factors for interpretation varied across the different factor selection criteria examined. Table 14 summarizes the results of each test and its suggested number of factors to retain. The parallel analysis and MAP tests were considered most heavily, given their reputation for greater accuracy in the literature (Osborne & Banjanovic, 2016). Given the exploratory nature of the EFA, the range of possible factor solutions suggested by all factor retention criteria were considered for interpretation. Previous factor analyses of the RBS-R with ASD samples resulted in four-, five- or six-factor solutions. Given the findings of various numbers of factors across existing studies, the wide range of results across the factor selection criteria roughly aligned with the current literature. Such results highlighted a need for exploration of the interpretability of a wide range of solutions in the present study. Across all selection criterion test results, examining factor solutions consisting of three to eight factors was suggested. However, more conservative criteria (i.e., Velicer's MAP and parallel analysis) suggested a range of factor solutions consisting of between four and six factors.

Table 14. Summary of Factor Retention Criterion Tests

Criterion	Suggested Number of Factors to Retain
Guttman-Kaiser Criterion	6, 8
Scree Test	3, 6
Velicer's MAP Test	4, 6
Parallel Analysis	4

Research Question 2

When special education staff rate students with ASD, how many factors should be retained to derive the most interpretable factor solution for the RBS-R?

Hypothesis 2a. *It is predicted that the most interpretable EFA solution will consist of multiple factors (i.e., $\geq$ two factors).*

Hypothesis 2b. *It is predicted that, within the most interpretable EFA solution, a Self-Injurious Behaviors factor will be present.*

Hypothesis 2c. *It is predicted that, within the most interpretable EFA solution, a Compulsive Behaviors factor will be present.*

All three hypotheses (2a, 2b, and 2c) were tested by examining the pattern and structure matrices, following the direct oblimin rotation (Jennrich & Sampson, 1966), to examine the factors in each factor solution for interpretability. For this study, suggestions stemming from Velicer's MAP and parallel analysis criteria, which suggested between four and six factors, were prioritized. In addition, factor solutions consisting of three factors or seven factors (i.e., considering each factor selection criterion +/- one factor) were also examined for interpretability to consider the inherent random error (e.g., sampling variation) that impacts factor selection criterion estimates. Given the support from one study (Georgiades et al., 2010) that suggested a two-factor solution for the Greek version of the RBS-R, a two-factor solution was also included in examination for interpretability. Therefore, given the exploratory nature of the analysis, a broad range of solutions was examined (i.e., all factor solutions consisting of between two and seven factors were examined for interpretability). Pattern matrices for the two-, three-, four-, six-, and seven-factor models can be found in Appendix B.

Interpretation. After factors were extracted and rotated, each potential factor solution was analyzed and provided with provisional names to determine the most interpretable factor solution. A team of four qualified researchers familiar with ASD and measurement issues independently analyzed all factor solutions. Following their independent interpretations, the team of researchers met to discuss the factor solutions and assess interpretive agreement and consensus. Initially, three of the six factor solutions were determined to be stronger than the others, based on clinical interpretability, item to factor structure, and number of cross-loadings. These three solutions (i.e., the four-, five-, and six-factor solutions) were considered to encompass similar concepts—wherein each successive solution appeared divide a prior factor into two, while retaining the integrity of the other factors in the prior solution. An overview of the six different factor solutions is provided in the text that follows.

Within the two-factor solution, the factors that made up the overall scale mirrored the early theories of repetitive behavior in the RRBI literature – including a higher- and lower-order distinction (Cuccaro et al., 2003; Georgiades et al., 2010; Szatmari et al. 2006). The first factor included items that reflected an insistence on sameness, rituals, and checking behaviors, while the second factor was defined primarily by items related to self-injury. The two-factor solution included substantive cross-loadings for 6 out of the 43 items and presented challenges in assigning credible factor names due to the broad range of item content within each factor that appeared to mask potentially important conceptual distinctions.

The three-factor solution included the first factor in the two-factor solution, emulating a category of insistence on sameness and checking behaviors. However, the second factor from the two-factor solution was observed to split into two different divisions of the original “self-injurious behavior” subscale. That is, the second factor in the three-factor solution included items representing more stereotyped motor behaviors (e.g., head rolling, while body rocking, hand, and finger mannerisms, etc.), while the third factor was defined primarily by items more reflective of self-injurious behavior (e.g., skin picking, scratching the skin, hitting self, etc.). However, the three-factor solution included substantive cross-loadings for 12 out of 43 items, making the item assignment to factors more problematic, as so many items appeared to be assessing multiple concepts. Furthermore, this solution included one item (item 22) with a non-substantive primary factor loading.

Similar to the pattern of emerging factors from the two-factor to the three-factor solution, the four-factor solution involved the splitting of a prior broader factor, while continuing to include two other factors that were largely intact from the preceding three-factor solution. That is, the four-factor solution continued to maintain the three-factor solution's distinction between self-injurious behavior and motor stereotypies. However, this solution also split the three-factor solution's first factor (i.e., insistence on sameness and checking behaviors) into two separate factors, now statistically supporting a conceptual distinction between "insistence on sameness" and "compulsive behaviors". This solution included one item (item 36) with a non-substantive primary factor loading. This solution also included several (11 total) items that were cross loaded on different factors, leaving room for questions about item uniqueness as part of factor assignments for this solution. Of note, later factor solutions (e.g., five- and six-factor solutions) made finer distinctions than the four-factor solution regarding item to factor conceptualization.

The five-factor solution highly regarded as the most interpretable and meaningful solution across all independent raters. This solution kept a similar structure to what was observed in the four-factor solution (maintaining the general integrity and interpretations of three of the four prior factors), but also divided the larger self-injurious behavior factor established in the four-factor solution into two conceptually meaningful factors. This division included a factor reflecting a construct related to Body-Focused Repetitive Behaviors – including behaviors such as skin picking, pulling hair or skin, and rubbing or scratching body parts. These items could readily be viewed as more intense or exaggerated versions of basic grooming behaviors. The other factor included more severe, overt forms of repetitive behavior that were more inherently damaging or physically dangerous to oneself, including hitting self with objects, hitting self against surfaces, hitting self with body parts, and biting self. Overall, the five factors were named Insistence on Sameness, Compulsive Behaviors, Stereotyped Behaviors, Body-Focused Repetitive Behaviors, and Self-Injurious Behaviors. This factor solution was favored for its clear presentation of conceptually and clinically meaningful factors, minimal item cross-loadings, and parsimony – when compared to the six-factor solution.

The six-factor solution was another relatively strong solution but presented more limitations than the five-factor solution. This factor solution included factors reflecting conceptual domain areas such as Sameness Behaviors, Compulsive Behaviors, Motor Stereotypies, Sensory-Seeking Behaviors, Self-Injurious Behaviors, and Restricted Behaviors. Like other solutions, there were several items cross-loaded onto multiple factors. One item (41: "Strongly attached to one specific object"), showed a loading on three factors, as opposed to just two factors like the other cross-loaded items. It was also noted that only three items loaded primarily onto Factor VI (i.e., Restricted Behavior factor), with one of them cross-loading substantively on another factor as well. Factor loadings for the three items on this sixth factor were only moderate to weak (0.46, 0.44, and 0.31), making it challenging to cleanly and appropriately interpret. Overall, given these limitations, especially regarding the sixth factor, the six-factor solution was not selected.

The seven-factor solution included multiple items that lacked unique loadings on any of the potential factors and yielded a factor (e.g., Factor VII) with only two items assigned to it. Factor constructs that emerged from this solution included Compulsive Behaviors, Insistence on Sameness, Stereotyped Behavior, Body-Focused Repetitive Behaviors, Restricted Behavior, Self-Injurious Behaviors, and Routinized Behaviors. The seventh factor (Routinized Behaviors) consisted of only two items, each with moderate primary loadings (.49 and .56), and both items substantively cross-loading on Factor I (Compulsive Behaviors). Though roughly nameable as a factor, based on its primary loadings, these substantive limitations undermined any unique, distinct interpretation of Factor VII. However, it is noteworthy that the other six factors that emerged were readily interpretable. This solution was not chosen due to the lack of a clear and distinct seventh factor.

The Retained, Most Interpretable Solution: Five-Factor Model. After concluding that the five-factor solution was the most interpretable, the four researchers discussed the factor names that they had each independently generated, and quickly reached consensus on the name and interpretation of each factor in the solution. Although all factors were generally named in alignment with each other across researchers, with most having the same name across researchers, Factor IV (ultimately named Body-Focused Repetitive Behaviors), present some difficulty. The item content made sense conceptually across researchers, but finding an all-encompassing name was challenging, given that this appeared to be a domain that had not been previously encountered as a distinct factor in the RBS-R literature. Overall, following discussion and unanimous consent, the researchers agreed on the assignment of the following factor names: (Factor I) Insistence on Sameness, (Factor II) Compulsive Behaviors, (Factor III) Stereotyped Behaviors, (Factor IV) Body-Focused Repetitive Behaviors, and (Factor V) Self-Injurious Behaviors. Table 15, below, shows the pattern matrix from study one's retained five-factor solution. This table includes all 43 items of the scale, with items sorted by factor and arranged from highest to lowest primary factor loading. (Appendix B includes the pattern matrices for the other factor solutions considered for interpretation.)

Table 15. Study One Five-Factor Solution Pattern Matrix

		Factor				
Item	Item Content	1	2	3	4	5
39	Insists that specific things take place at specific times	**<u>0.96</u>**	0.12	0.03	0.02	0.07

Table 15 (cont'd)

38	Insists on same routine, household, school, or work schedule everyday	**<u>0.83</u>**	0.07	0.10	0.06	0.10
33	Insists on sitting at the same place	**<u>0.72</u>**	0.14	0.14	0.12	0.09
34	Dislikes changes in appearance or behavior of the people around him/her	**<u>0.65</u>**	0.08	0.09	**0.48**	0.09
37	Resists changing activities; Difficulty with transitions	**<u>0.64</u>**	0.04	0.10	0.12	0.15
30	Objects to visiting new places	**<u>0.61</u>**	0.06	0.00	0.16	0.13
28	COMMUNICATION/SOCIAL INTERACTIONS (Repeats same topic(s) during social interactions; Repetitive questioning; Insists on certain topics of conversation; Insists that others say certain things or respond in certain ways during interactions)	**<u>0.58</u>**	0.01	0.30	0.14	0.30
35	Insists on using a particular door	**<u>0.55</u>**	0.25	0.12	0.24	0.04
41	Strongly attached to one specific object	**<u>0.53</u>**	0.04	**0.34**	0.06	0.11
27	PLAY/LEISURE (Insists on certain play activities; Follows a rigid routine during play/leisure; Insists that certain items be present/available during play/leisure; Insists that other persons do certain things during play	**<u>0.51</u>**	0.04	0.21	0.23	0.21
40	Fascination, preoccupation with one subject or activity (e.g., trains, computers, weather, dinosaurs)	**<u>0.45</u>**	0.15	0.26	0.01	0.10
32	Insists on walking in a particular pattern (e.g., straight line)	**<u>0.43</u>**	0.09	0.06	0.19	0.07
26	TRAVEL/TRANSPORTATION (insists on taking certain routes/paths; Must sit in specific location in vehicles; Insists that certain items be present during travel, e.g., toy or material; Insists on seeing or touching certain things or places during travel such as a sign or store)	**<u>0.43</u>**	0.24	0.02	0.12	0.10

Table 15 (cont'd)

31	Becomes upset if interrupted in what he/she is doing	**0.43**	0.03	0.17	0.16	**0.31**
20	HOARDING/SAVING (Collects, hoards, or hides specific items)	**0.32**	0.19	0.04	0.08	0.29
36	Likes the same CD, tape, record, or piece of music played continually; Likes same move/video or part of movie/video	**0.31**	0.22	0.26	0.17	0.14
17	WASHING/CLEANING (Excessively cleans certain body parts; Picks at lint or loose threads)	0.25	**0.88**	0.12	0.24	0.20
16	COMPLETENESS (Must have doors opened or closed; Takes all items out of a container or area)	0.01	**0.86**	0.04	0.00	0.00
15	ARRANGING/ORDERING (arranges certain objects in a particular pattern or place; Need for things to be even or symmetrical)	0.06	**0.82**	0.13	0.05	0.17
18	CHECKING (Repeatedly checks doors, windows, drawers, appliances, clocks, locks, etc.)	0.28	**0.72**	0.20	0.03	0.01
19	COUNTING (Counts items or objects; Counts to a certain number or in a certain way)	0.03	**0.72**	0.17	0.05	0.03
21	REPEATING (Need to repeat routine events; In/outdoor, up/down from chair, clothing on/off)	0.22	**0.70**	0.09	0.05	0.23
25	SELF-CARE - BATHROOM AND DRESSING (Insists on specific order of activities or tasks related to using the bathroom, to watching, showering, bathing, or dressing; Arranges items in a certain way in the bathroom or insists that bathroom items not be moved; Insists on wearing certain clothing items)	0.26	**0.61**	0.03	0.06	0.10
29	Insists that things remain in the same place(s) (e.g., toys, supplies, furniture, pictures, etc.)	**0.45**	**0.54**	0.24	0.11	0.13
23	EATING/MEALTIME (Strongly prefers/insists on eating/drinking only certain things; Eats or drinks items in a set order; Insists that meal related items are arranged in a certain way)	0.24	**0.43**	0.20	0.26	0.01

Table 15 (cont'd)

24	SLEEPING/BEDTIME (Insists on certain pre-bedtime routines; Arranges items in room "just so" prior to bedtime; Insists that certain items be present with him/her during sleep; Insists that another person be present prior to or during sleep)	**0.36**	<u>**0.42**</u>	0.13	0.30	0.08
2	HEAD (rolls head, nods head, turns head)	0.14	0.10	<u>**0.79**</u>	0.12	0.09
1	WHOLE BODY (Body rocking, body swaying)	0.08	0.06	<u>**0.76**</u>	0.04	0.07
4	LOCOMOTION (turns in circles, whirls, jumps, bounces)	0.13	0.08	<u>**0.74**</u>	0.09	0.11
5	OBJECT USAGE (spins or twirls objects, twiddles or slaps or throws objects, lets objects fall out of hands)	0.07	0.01	<u>**0.69**</u>	0.00	0.22
3	HAND/FINGER (Flaps hands, wiggles or flicks fingers, claps hands, waves or shakes hand or arm)	0.06	0.06	<u>**0.66**</u>	0.08	0.05
6	SENSORY (covers eyes, looks closely or gazes at hands or objects, covers ears, smells, or sniffs items, rubs surfaces)	0.09	0.09	<u>**0.63**</u>	0.24	0.04
43	Fascination, preoccupation with movement/things that move (e.g., fans, clocks)	0.23	0.14	<u>**0.50**</u>	0.18	0.14
42	Preoccupation with part(s) of object rather than the whole object (e.g., buttons on clothes, wheels on cars)	0.30	0.12	<u>**0.48**</u>	0.06	0.12
22	TOUCH/TAP (Need to touch, tap, or rub items, surfaces, or people)	0.00	0.28	<u>**0.38**</u>	0.13	0.03
13	INSERTS FINGER OR OBJECT (eye-poking, ear-poking)	0.09	0.06	0.14	<u>**0.72**</u>	0.01
14	SKIN PICKING (picks at skin on face, hands, arms, legs, or torso)	0.19	0.24	0.11	<u>**0.68**</u>	0.16
11	PULLS (pulls hair or skin)	0.13	0.02	0.05	<u>**0.65**</u>	0.26

Table 15 (cont'd)

12	RUBS OR SCRATCHES SELF (Rubs or scratches marks on arms, leg, face, or torso)	0.01	0.08	0.25	**<u>0.63</u>**	0.24
9	HITS SELF WITH OBJECT (Hits or bangs head or other body area with objects)	0.01	0.11	0.01	0.14	**<u>0.83</u>**
8	HITS SELF AGAINST SURFACE OR OBJECT (hits or bangs head or other body part on table, floor, or another surface)	0.02	0.14	0.14	0.12	**<u>0.81</u>**
7	HITS SELF WITH BODY PART (Hits or slaps head, face, or other body area)	0.10	0.03	0.45	0.27	**<u>0.76</u>**
10	BITES SELF (Bites hand, wrist, arm, lips, or tongue	0.13	0.15	0.19	0.23	**<u>0.43</u>**

Note. The primary factor loading for each item (reflecting the factor to which it is assigned) is underlined, while all loadings > .30 are bolded.

Factor I: Insistence on Sameness. The first factor in the selected five-factor solution consisted of 16 items with the highest loading items clearly reflecting an insistence on sameness. The highest loadings on this factor were item 39 (.96) related to an insistence that things take place at a specific time, item 38 (.83) related to insistence on the same daily routines, and item 33 (.72) relating to insistence on sitting at the same place. Two items had cross-loadings (i.e., loadings > .30 on a different factor included in the solution). This was item 41 (Factor III), which refers to an individual's attachment to specific objects (.34). Additionally, item 34 referring to a dislike in a change of others' appearances, had a cross-loading of .48 with Factor IV (Body-Focused Repetitive Behaviors).

Factor II: Compulsive Behaviors. The second factor, related to compulsive and ritualistic behaviors, was composed of 10 items. The items with the highest loadings were item 17 (.88; excessive washing and cleaning oneself), item 16 (.86; completeness – insisting on having doors open/closed and empty containers), and item 15 (.82; arranging and ordering). Two items had cross-loadings, both with Factor I (Insistence on Sameness), including item 29 (.45; insisting that things remain in the same place), and item 24 (.36; related to insisting on certain bedtime routines and arranging items "just so" before bedtime).

Factor III: Stereotyped Behaviors. The third factor, related to body/motor-focused stereotyped behavior, was comprised of nine items. Items with the highest loadings on this factor included item 2 (.79; head rolling, turning, and nodding); item 1 (.76; whole-body movements); and item 4 (.74); turning in circles, jumping, and bouncing). There were no cross-loadings for items on this factor.

Factor IV: Body-Focused Repetitive Behaviors. The fourth factor consisted of four items (items 11-14), related to repetitive behaviors that appeared to be associated with, or exaggerations of, self-grooming behaviors. The items in this factor all pertained to an aspect of repetitive behaviors that in more extreme forms are commonly associated with self-injurious behaviors (e.g., eye poking or gauging) or with certain compulsive behaviors (e.g., hair pulling observed in trichotillomania)—though these behaviors when in their mild form may not necessarily be considered pathological. It is noteworthy that the four items from this factor were part of the RBS-R's originally conceived Self-Injurious Behavior subscale, and they effectively split off from the other more inherently violent self-injury items when moving from the four-factor to the five-factor solution in the present EFA. More specifically, the items that constituted this factor included item 13 (.72; eye poking or ear poking), item 14 (.68; skin picking), and item 11 (.65; hair pulling). The final item, item 12, included information related to rubbing or scratching oneself and had an item loading of .63. There were no cross-loadings for the four items that loaded primarily on this factor.

Factor V: Self-Injurious Behaviors. The fifth and final factor was related to more inherently violent behaviors (e.g., hitting, biting) that are most connected with injury to oneself. This factor consisted of four items (items 7-10). It is noteworthy that these items were part of the RBS-R's original Self-Injurious Behavior subscale. The three highest loading items on this factor were item nine (.83; hitting the self with objects/head banging), item eight (.81; hitting body parts on surfaces), and item seven (.76; hitting the self with another body part). The final item, item 10, was related to biting oneself and had a weaker item loading (.43) when compared to the other items in this factor. There were no cross-loadings for any of the four items that loaded primarily on this factor.

Overall, the five-factor solutions included the fewest number of items with cross-loadings, compared to the other solutions yielded from the EFA. The items on the five-factor solution that yielded cross-loadings included items 34 (Factor I), 41 (Factor I), 31 (Factor I), 29 (Factor II), 23 (Factor II), and 34 (Factor II). Additionally, some items produced weaker loadings on their assigned factors. For example, items 20 and 36 (both Factor I) yielded loadings of .32 and .31, respectively. These two items included loadings on other factors that were not considered adequate for assignment or cross-loadings (e.g., .29 or .26). Interpretation of these findings is included in chapter four.

Once the five-factor solution was interpreted fully, Hypotheses 2a, 2b, and 2c were assessed. First, hypothesis 2a, which predicted that the most interpretable EFA solution will consist of multiple factors (i.e., > two factors), was supported, as a five-factor solution was selected as the most interpretable and meaningful solution. Second, hypothesis 2b, which predicted that a Self-Injurious Behavior factor would be present in the most interpretable EFA factor solution, was supported, as a clear Self-Injurious Behavior factor (i.e., Factor V) did emerge—consisting of items that were part of the RBS-R's originally-conceived Self-Injurious Behavior subscale. Finally, hypothesis 2c, which predicted that a Compulsive Behaviors factor will be present in the most interpretable EFA solution, was supported, as a clear Compulsive Behaviors factor (i.e., Factor II) did emerge.

Research Question 3

When special education staff rate students with ASD, does the most interpretable factor structure yield substantive correlations amongst at least some of the factors of the RBS-R?

Factor solutions subject to oblique rotation (i.e., direct oblimin) as part of the EFA in order to determine whether there were substantive correlations between at least some of the RBS-R factors. Table 16 below displays the inter-factor correlations for the selected five-factor solution.

Table 16. Study One Five-Factor Solution Inter-Factor Correlation Matrix

Factor Number and Name	Factor Number				
	1	2	3	4	5
Factor I: Insistence on Sameness	1.00				
Factor II: Compulsive Behaviors	**.56**	1.00			
Factor III: Stereotyped Behaviors	**.41**	.26	1.00		
Factor IV: Body-Focused Repetitive Behaviors	.10	.16	.22	1.00	
Factor V: Self-Injurious Behaviors	**.43**	.27	**47**	.28	1.0 0

Note. Correlations greater than or equal to .30 are bolded.

Four of the factors were each substantively correlated (i.e., > .30; Beavers et al., 2013) with at least one other factor. In addition, substantive inter-factor correlations were assessed using Cohen’s (1988) effect size standards for correlation coefficients, with correlations $\geq$ .50

considered to be large and correlations < .50 and $\geq$ .30 as medium or moderate. Factor I (Insistence on Sameness) had a large correlation with Factor II (Compulsive Behaviors: .56), as well as moderate correlations with Factor III (Stereotyped Behaviors) and Factor V (Self-Injurious Behaviors). Factor V (Self-Injurious Behaviors) was also moderately correlated with Factor III (Stereotyped Behaviors: .47). Thus, to respond clearly to research question three, one large and three moderate or medium substantive inter-factor correlations were present. Therefore, the obliquely rotated (correlated) factor solution was retained.

Internal Consistency Reliability

Though not central to a specific research question, internal consistency estimates were calculated and reported here for the sake of thoroughness when considering using the factors to create clearly interpretable RBS-R subscales. Internal consistency reliability estimates were calculated using both ordinal alpha and Cronbach's alpha. Ordinal alpha was considered the primary internal consistency reliability estimate due to the ordinal nature of the item data and use of the polychoric correlation matrix. However, both ordinal and Cronbach's alpha estimates were reported in order to maintain a common standard for comparison with previous studies that reported only Cronbach's alpha estimates. Table 17 includes the internal consistency reliability estimates for the selected five-factor solution.

Table 17. Internal Consistency Reliability Estimates

Factor	Name	Number of Items	Cronbach's Alpha	Ordinal Alpha
1	Insistence on Sameness	16	.873	.934
2	Compulsive Behaviors	10	.867	.937
3	Stereotyped Behaviors	9	.854	.901
4	Body-Focused Repetitive Behaviors	4	.720	.871
5	Self-Injurious Behaviors	4	.804	.893

Ordinal alpha estimates ranged from .871 to .937 with three out of the five factors $\geq$ .90 and all five factors $\geq$ .80. Cronbach's alpha estimates ranged from .720 to .873 with four out of the five factors $\geq$ .80. According to Salvia et al. (2017), scales with reliability estimates $\geq$.70 are adequate for weekly progress monitoring, $\geq$ .80 for screening, and $\geq$ .90 for decision making. Also, Nunnally (1978) stated that scales should have reliability estimates above .70 for research purposes—a standard that was met for all factors regardless of alpha coefficient type. Finally, under the internal consistency criteria set by Murphy and Davidshofer (1998**),** ordinal alpha estimates for three of the five factors (scales) fell in the excellent range (i.e., .90 to .99) and two fell in the moderately high or good range (i.e., .80 to .89). Four of the five Cronbach's alpha values fell in the moderately high or good range, leaving one value (for Factor 4) in the acceptable range (i.e., .70 to .79).

Research Question 4

When special education staff rate students with ASD, to what extent does the six-factor solution from the present EFA (regardless of whether this is the retained solution) correspond to the six-subscale structure originally proposed by Bodfish et al. (2000)?

Overall, there were some notable similarities between the study one EFA six-factor solution and the published Bodfish et al. (2000) six-factor solution. Namely, five of the six-factor names in the study one EFA solution are also present in the Bodfish et al. (2000) six-factor solution used to create the original six subscales of the RBS-R. Again, the current EFA factor names were based on input from the independent researchers who evaluated and interpreted the presented factor solutions. Thus, it is noteworthy that they converged on the same or similar factor names for those five factors. However, it is very important to recognize that, in the EFA, only 25 out of the RBS-R's 43 items (i.e., 58.14% overlap) were assigned (based on primary loadings) to the same factor or conceptual subscale across the two models. A split between the original solution's Self-Injurious Behaviors factor in the present EFA's solution led to a shift in several items on factors. Across both models, no items were left out of the factor solution, which allowed for a common and comprehensive set of item comparisons across solutions.

It is also noted that the present study's model yielded higher factor loadings for most of the items compared to the published RBS-R solution. Information regarding the factor loadings for the published RBS-R was sourced from Lam (2004), where results were stated to come from an unpublished presentation by Bodfish and Lewis (2002), involving a principal components analysis, as described in Chapters 2 and 3. These item-to-factor loadings were stated to come from Lam's personal communication with Bodfish and are reported in the Lam (2004) dissertation. The item loadings have been used to support the assumption of considerable structure for the Bodfish et al. (2000) RBS-R but have not been used to assess the RBS-R's validity by the test authors. See Table 18 for a comparison of the factor names between models. Table 18 also includes factor-loading comparisons for each item across the two solutions. Overall, the six-factor solution from the present study's EFA held various similarities to the Bodfish et al. (2000) model, with major differences highlighted in the number of factors and item loading on factors. See Table 18 below for a comparison of factor names for the Bodfish et al. (2000) six-factor solution and the six-factor solution that was generated (though not ultimately chosen) from the EFA in this study.

The Insistence on Sameness (current EFA)/Sameness Behaviors (Bodfish et al., 2000) subscales, Stereotyped Behaviors subscales, and Self-Injurious Behaviors subscales in both factor solutions were very similar in terms of their highest loading items and items present in each factor. Of note, the Self-Injurious Behaviors subscale in the present study differed from that found in the Bodfish et al. (2000) model due to the splitting up of Bodfish et al.'s SIB factor into two distinct factors in the present study EFA. That is, the SIB-focused items on the original, published solution include items 7 to 14, while items 7-10 are included in the present study's Self-Injurious Behaviors factor. In this study, items 11-14 were assigned to their own factor, named Body-Focused Repetitive Behaviors--based on the apparent conceptual scope of the item content in this domain. The highest loadings in the Restricted Behaviors subscale varied between the present study's factor solution and the published solution, as the solution from the present EFA assigned most of the items from this factor in the original solution to other factors. Further, two of the three items in the EFA's Restricted Interest factor were originally present in Bodfish et al. (2000)'s Sameness Behavior factor/subscale.

Table 18. Comparison of Factors and Item Content in the Present EFA's Six-Factor Solution (Not Retained) to the Original Six Factors of the RBS-R (Bodfish et al., 2000)

Present Study			Published Bodfish et al. (2000) Study*		
Highest to Lowest Item Loadings in Present Study			Highest to Lowest Item Loadings in Bodfish et al. (2000) model, as cited in Lam (2004)		
Factor Name	Item	Loading	Factor Name	Item	Loading
Insistence on Sameness	39	0.98	**Sameness Behavior**	39	.85
	38	0.90		38	.84
	37	0.67		37	.62
	33	0.64		35	.51
	30	0.60		30	.40
	26	0.57		33	.34
	34	0.54		31	.24
	28	0.50		34	.21
	24	0.49		36	.19
	27	0.48		29	.14
	32	0.46		32	.08
	23	0.42			
	31	0.42			
	41	0.39			
57% overlap *(8 out of 14 items)*[a]			*73% overlap (8 out of 11 items)*[b]		

Table 18 (cont'd)

Compulsive Behaviors	17	0.88	**Compulsive Behavior**	21	.71
	16	0.86		18	.47
	15	0.82		15	.43
	19	0.73		16	.32
	18	0.72		22	.31
	21	0.70		17	.24
	25	0.60		20	.20
	29	0.54		19	.18
75% overlap (6 out of 8 items)[a]			*75% overlap (6 out of 8 items)[b]*		
Stereotyped Behaviors	2	0.79	**Stereotyped Behavior**	3	.68
	1	0.75		5	.63
	4	0.74		4	.60
	5	0.67		1	.50
	3	0.65		2	.44
	6	0.64		6	.41
	42	0.44			
	43	0.46			
	22	0.36			
67% overlap (6 out of 9 items)[a]			*100% overlap (6 out of 6 items)[b]*		
Self-Injurious Behaviors	9	0.83	**Self-Injurious Behavior**	7	.79
	8	0.82		9	.65
	7	0.75		8	.63
	10	0.43		10	.56
	20	0.31		11	.32
				12	.22
				13	.08
				14	.05
80% overlap (4 out of 5 items)[a]			*50% overlap (4 out of 8 items)[b]*		
Body-Focused Repetitive Behaviors	13	0.70	**Ritualistic Behavior**	27	.60
	14	0.69		25	.59
	11	0.67		24	.45
	12	0.63		23	.32
				28	.26
				26	.07
0% overlap (0 out of 4 items)[a]			*0% overlap (0 out of 6 items)[b]*		
Restricted Behaviors	40	0.46	**Restricted Behavior**	40	.53
	35	0.44		41	.47
	36	0.31		42	.01
				43	.01
33% overlap (1 out of 3 items)[a]			*25% overlap (1 out of 4 items)[b]*		

Table 18 (cont'd)

Note. This table provides the percentage of overlapping items between the factors from the published RBS-R (Bodfish et al., 2000), which stemmed from an unpublished presentation (Bodfish & Lewis, 2002) and the study one EFA six-factor solution (not retained).

[a]Overlap percentage represents the amount of item overlap for the study one EFA factors, using the total number of items in the factor from the present EFA as the denominator.

[b]Overlap percentage represents the amount of item overlap for the Bodfish et al. (2000) factors, using the total number of items in the factor/RBS-R subscale from the results of the Bodfish et al. (2000) factor analysis.

Eight items overlapped between the published Bodfish et al. (2000) Sameness Behavior factor, which included 11 items (73% overlap), and between the present study's Insistence on Sameness factor, which included 14 items (57% overlap). The differences in overlap percentage are present due to the different numbers (i.e., 11 or 14 items, in this case) used in the denominator. The EFA's Insistence on Sameness factor included six additional items (items 23, 24, 26, 27, 28, and 41, past the three (items 29, 35, and 36) that were not included in the factor. Six items from the original Bodfish et al. (2000) Compulsive Behavior factor (eight items; 75% overlap) and present study EFA's Compulsive Behaviors factor (also 8 items; 75% overlap) occurred. However, the present study's Compulsive Behaviors factor included items 25 and 29 (originally on the published RBS-R's "Ritualistic Behaviors" subscale) and the original Bodfish et al. (2000) solution excluded items 20 and 22 that the present study inserted. All six items (100% overlap) from the published Bodfish et al. (2000)'s factor solution for Stereotyped Behaviors were included in the present study's same-named factor. The present study added three additional items, including item 22 (originally assigned to the Compulsive Behavior subscale) and items 42 and 43 (originally assigned to the Restricted Behavior subscale), creating a 67% overlap of items when nine are included in the denominator.

Four out of eight items from the Bodfish et al. (2000) Self-Injurious Behavior subscale (eight items; 50% overlap) and the present study's Self-Injurious Behaviors subscale (five items; 80% overlap) overlapped between the two individual factor solutions. The present study's SIB factor included the addition of item 20 (hoarding and saving, originally assigned to the Compulsive Behavior subscale for Bodfish et al, 2000). The most notable difference is that the original RBS-R's SIB factor was split, creating a second factor in the present study, labeled Body-Focused Repetitive Behaviors. Because this factor did not exist in the published model, it cannot be meaningfully compared to another distinct factor from the published model. Similarly, the Restricted Behaviors factor from the published solution was dissolved into other factors in the present study EFA. Thus, it could not be meaningfully compared to another distinct factor in the published model. One item overlapped between the published Bodfish et al. (2000) Restricted Interests subscale (three items; 33% overlap) and the study one EFA Restricted Interests subscale (four items, 25% overlap). The only item to overlap was item 40, which refers to a fascination or preoccupation with a particular subject or activity.

Research Question 5

When special education staff rate students with ASD, to what extent does the five-factor solution from the present EFA (regardless of whether it is the retained solution) correspond to the five-factor solution reported by Lam and Aman (2007)?

Table 19 compares factor names, item content, and item loading on factors for the present study's EFA and the Lam and Aman (2007) factor solutions. Both solutions yielded similar factor names, apart from one factor from each solution (Body-Focused Repetitive Behaviors in the present study and Restricted Interests in the Lam and Aman [2007] solution). Further, identical names were assigned for Compulsive Behaviors, Stereotyped Behaviors, and Self-Injurious Behaviors, while an additional factor (Insistence on Sameness/Sameness Behaviors) suggested a similar domain of sameness behaviors. Further, the highest loading items on the Sameness Behaviors/Insistence on Sameness factors were identical (items 39 and 38). Similarly, the highest loading item for the Self-Injurious Behaviors factors was the same (item 9) in both solutions as well.

Table 19. Comparison of Factors and Item Content in the Present EFA's Five-Factor Solution to the Five-Factor Lam and Aman (2007) Solution

Present Study			Published Lam and Aman (2007)		
Highest to Lowest Item Loadings in Present Study			Highest to Lowest Item Loadings in Lam and Aman Study		
Factor Name	Item	Loading	Factor Name	Item	Loading
Insistence on Sameness	39	0.96	**Sameness Behavior**	38	.78
	38	0.83		39	.77
	33	0.72		37	.74
	34	0.65		30	.71
	37	0.64		33	.63
	30	0.61		31	.57
	28	0.58		32	.57
	32	0.57		34	.53
	26	0.56		27	.51
	35	0.55		26	.48
	41	0.53		35	.48
	27	0.51		29	.41
	40	0.45		36	.38
	31	0.43		28	.37
	20	0.32			
	36	0.31			
75% overlap (12 out of 16 items)[a]			*86% overlap (12 out of 14 items)[b]*		
Compulsive Behaviors	17	0.88	**Compulsive Behavior**	15	.73
	16	0.86		18	.58
	15	0.82		17	.55
	18	0.72		16	.55
	19	0.72		19	.52
	21	0.70		20	.47
	25	0.61			
	29	0.54			

Table 19 (cont'd)

	23	0.43			
	24	0.42			
50% overlap (5 out of 10 items)[a]			*83% overlap (5 out of 6 items)[b]*		
Stereotyped Behaviors	2	0.79	**Stereotyped Behavior**	3	.65
	1	0.76		5	.63
	4	0.74		4	.57
	5	0.69		2	.51
	3	0.66		6	.49
	6	0.63		1	.47
	43	0.50		43	.46
	42	0.48		42	.44
	22	0.38		22	.35
100% overlap (9 out of 9 items)[a]			*100% overlap (9 out of 9 items)[b]*		
Self-Injurious Behaviors	9	0.83	**Self-Injurious Behavior**	9	.70
	8	0.81		12	.69
	7	0.76		7	.67
	10	0.43		11	.62
				14	.57
				8	.55
				13	.51
				10	.51
100% overlap (4 out of 4 items)[a]			*50% overlap (4 out of 8 items)[b]*		
Body-Focused Repetitive Behaviors	13	0.72	**Restricted Behavior**	40	.65
	14	0.68		36	.50
	11	0.65		41	.45
	12	0.63		40	.65
				36	.50
				41	.45
0% overlap (0 out of 4 items)[a]			*0 % overlap (0 out of 3 items)[b]*		

Note. This table provides the percentage of overlapping items between the factors from the Lam and Aman (2007) five-factor solution and the retained study one EFA five-factor solution. Item loadings from the Lam and Aman (2007) column of the table are taken from Lam (2004; pg. 43, Table 3).

[a]Overlap percentage represents the amount of item overlap for the study one EFA factors, using the total number of items in the factor from the present EFA as the denominator.

Table 19 (cont'd)

[b]Overlap percentage represents the amount of item overlap for the Lam and Aman (2007) factors, using the total number of items in the factor/RBS-R subscale from the Lam and Aman (2007) factor analysis results.

Twelve items overlapped between the original Lam and Aman (2007) Sameness Behavior factor, which included 14 items (86% overlap), and between the present study's Insistence on Sameness factor, which included 16 items (75% overlap). The differences in overlap percentage are present due to the different numbers (i.e., 14 or 16 items) used in the denominator. The present study's Insistence on Sameness factor included three additional items (items 20, 40, and 41) that were not included in Lam and Aman (2007)'s solution. Five items overlapped between the original Lam and Aman (2007) Compulsive Behaviors factor (six items; 83%) and between the study one EFA 's Compulsive Behaviors factor (10 items 50% overlap).

Item 20 was the only item not included in the present study, while items 21, 23, 24, 25, and 29 were added in. All these listed additional items were excluded from the Lam and Aman (2007) factor solution as they did not load strongly enough onto one particular factor (Lam, 2004). All items (nine out of nine – 100% overlap) from the Stereotyped Behaviors subscales in both factor solutions overlapped. Though the order of highest to lowest items on each factor differed slightly, there were no additions or substitutions of items across the two factors. Similar to the Self-Injurious Behaviors factor described for the EFA's six-factor model, there was a division of the scale's items relating to self-injury. Again, only four out of the eight original items were included in the present study's factor, as the rest of the items were assigned to a new factor, Body-Focused Repetitive Behaviors. Therefore, no items overlapped for the Body-

Focused Repetitive Behaviors subscale (0% overlap) from the present study, and no items overlapped for the Restricted Interests subscale from Lam and Aman (2007).

Study Two: CFA

The study two CFAs utilized a separate validation sample from the exploratory sample used for the EFA in study one. The validation sample for the CFA included 233 individuals aged three to 22 years old with ASD. Like study one, the sample included the same types of raters – consisting of special education staff (e.g., special education teachers, teaching assistants, speech/language pathologists, physical therapists, occupational therapists, behavior technicians, individual student aides, whole-classroom aides, and trained volunteer assistants employed by the special education agency) who were each familiar with the participant they rated. (See Chapter 3 for more details regarding the sample for study two.)

Model Specification and Identification

Several CFA models were tested in the study-two CFA analyses. Models included a) the five-factor solution derived and selected as most interpretable in study one, b) the original Bodfish et al. (2000) six-factor model, b) the five-factor Mirenda et al. (2010) model, c) the five-factor Bishop et al. (2013) model, d) the six-factor Martínez-González and Piqueras (2017) model, e) the four-factor Russell et al. (2019) model, and f) the five-factor model from Sturm et al., (2022). Overall, the fit of six different CFA models was assessed. See Appendix C for the path diagrams of the tested CFA models, included in Figures C1 through C7.

Model Estimation and Fit

The polychoric correlation matrix based on the ordinal item data was used as the input for model estimation using Mplus Version 8.2 (Muthén & Muthén, 1998-2017). Due to the ordinal scaling and non-normal aspects of the RBS-R item data, the estimator used was Mplus'

Weighted Least Squares Mean Variance (WLSMV) approach, which was designed to accommodate such data. With this estimator, results yielded indices of model fit within the study two dataset. Fit indices included were the chi-square test (χ^2), standardized root mean square residual (SRMR), root mean square error of approximation (RMSEA), comparative fit index (CFI), and Tucker-Lewis index (TLI). Additional information indices were also included and were calculated using the robust maximum likelihood (MLR) estimator. These indices included the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). MLR was used to obtain these information indices because the WLSMV estimator does not allow for their calculation.

Multiple fit indices were generated for the determination of fit for each factor model and were used to compare the relative fit of the models that included the same number of items in their models. That is, direct comparisons of the Bodfish et al. (2000) model, the Mirenda et al. (2010) model, the Martínez-González and Piqueras (2017) model, the Sturm et al. (2022) model, and the model of the study one EFA for fit were allowed, as they all used the same number of items, and therefore same variance-covariance matrix for comparison. All other models (Bishop et al., 2013; Lam & Aman, 2007; Russell et al., 2019) formally discussed throughout studies one and two did not use the same number of items in their covariance matrix, as they omitted some items for various reasons. Therefore, these models were not comparable to the study one EFA five-factor model. Additionally, the Mplus DIFFTEST was intended to be used for model comparison in the case where a model was nested within another model. (The DIFFTEST is very useful in the model comparison context because it is a significance test.) However, this was not the case for any observed models, and, thus, the DIFFTEST was not used. Yet, for models that included all 43 items, it was still possible to compare relative model fit using the AIC and BIC indices generated through the MLR estimator—even when the models being compared were not nested.

Research Question 6

When special education staff rate students with ASD, does the model generated by the EFA of the RBS-R in the exploratory sample fit the inter-item covariance matrix for the validation sample?

Hypothesis 6. *It is predicted that the most interpretable model generated by the EFA in study one will adequately fit the inter-item covariance matrix of the CFA validation sample.*

Hypothesis 6 was addressed using the Mplus WLSMV estimator, using the WLSMV-adjusted χ^2, SRMR, RMSEA, CFI, and TLI. (For the sake of thoroughness, the seven other models reviewed from the RBS-R research literature were also evaluated for fit using these indices, and their results are reported below along with the results for the five-factor model derived from the EFA in study one.)

Results for all seven models examined across absolute fit indices are included below in Table 20. Absolute fit indices assess if the model's predicted variance-covariance matrix is the same as (equivalent to) the sample variance-covariance matrix (Harrington, 2009). Under the WLSMV-adjusted chi-square index, any statistically significant result ($p < .05$) suggests that the hypothesized model does not precisely fit the data. The chi-square statistic for the study one EFA five-factor model was statistically significant and therefore, strictly speaking, did not meet this criterion for an exact model fit. It is of note, however, that all seven additional models examined in this study also yielded statistically significant chi-square values ($p < .001$), and, thus, all failed to meet this criterion for model fit as well. Such results are common in the CFA literature, as the chi-square fit index very strictly assesses exact fit and its power to detect model discrepancies is highly dependent on sample size. Thus, larger sample sizes, typically desirable for purposes of stable estimates in a factor analysis or other structural equation modeling context, can too easily yield statistically significant results over minor discrepancies (Byrne, 2012). Because of this issue, a significant chi-square may not be a cause for substantive concern in this context. In fact, because of the prevalence of this issue in many research contexts, Alvi et al. (2020), based on work by Wheaton, et al. (1977) and Cole (1987), indicated that a descriptive ratio of the chi-square value divided by the degrees of freedom may be preferable and a result ≤ 2 would suggest better fit. Considering this ratio and interpretive guideline, the model based on the present EFA has a ratio of 1.571, which met this rule of thumb. Among the seven other models evaluated, three (i.e., Lam & Aman, 2007; Mirenda et al., 2010; Russell et al., 2019) failed to meet this criterion, while the four others (i.e., Bishop et al., 2013; Bodfish, 2000; Martínez-González & Piqueras, 2017; Sturm et al., 2022) yielded ratios < 2.

The SRMR was also used to determine absolute fit, which measures how different the hypothesized model is from a perfect fit of zero (i.e., higher SRMR values convey lack of fit). Hu and Bentler (1999) state that a cutoff value “close to .08” is recommended. The five-factor model from study one yielded an SRMR of 0.103, which was above the suggested cutoff. However, all models yielded an SRMR above this value, with the exception of the published RBS-R model by Bodfish et al. (2000), which attained an SRMR of exactly 0.08.

Table 20. CFA Model Results: Absolute Fit Indices

Model	c^2	*df*	*p*	SRMR
Bodfish (2000) six-factor model[a]	1205.892	845	< .001	0.080
Lam & Aman (2007) five- factor solution[b]	3900.072	855	< .001	0.209
Mirenda et al. (2010) five-factor model[a]	2532.236	850	< .001	0.137
Bishop et al. (2013) five-factor model[b]	1202.263	655	< .001	0.118
Martínez-González & Piqueras (2017) six-factor model[a]	1283.736	845	< .001	0.108

Table 20 (cont'd)

Russell et al. (2019) four- factor model[b]	4359.426	862	< .001	0.245
Sturm et al. (2022) six-factor model[a]	1302.698	850	< .001	0.103
Study one EFA five-factor model[a]	**1289.568**	**850**	**< .001**	**0.103**

Note. Only five of the listed models used the full 43-item set from the original RBS-R. These models are depicted below in superscript.

[a]Indicates models using the full 43 items from the original RBS-R.

[b]Indicates models using less than the full 43 items from the original RBS-R, omitting at least one or more items from the tested model.

Examination of all models using fit indices that assessed for parsimony was also considered in this study. Table 21 displays the RMSEA parsimony correction index and its confidence interval for each model. Parsimony correction indices are similar to absolute fit indices, except they take degrees of freedom (*df*) into account, which results in an increasing penalty for more complex models with fewer *df*. The RMSEA was used to measure the lack of fit relative to the population. Under RMSEA, a perfect fit is reflected by a value of 0. Values lower than .05 are considered to be a "close fit", values between .05 and .08 are a "reasonable fit", and values greater than or equal to .10 are not acceptable (Browne & Cudeck, 1992). Further, Hu and Bentler (1999) consider RMSEA values .06 and above as reasonable. The study-one five-factor model yielded a close fit (RMSEA = 0.047) and was one of four models examined within study two to fall in the "close fit" category. The other models to fit into this category were the Martínez-González and Piqueras (2017) six-factor model, the Sturm et al. (2022) five-factor model, and the Bodfish et al. (2000) original six-factor model. One of the models (Lam & Aman [2007] five-factor model) was considered reasonable (RMSEA = 0.060) but exceeded the cutoff for a close fit. The remaining studies (Bishop et al. [2013] five-factor model; Mirenda et al.

[2010] five-factor model, Russell et al. [2019] four-factor model) yielded higher values in the "not acceptable" range.

Table 21. CFA Model Results: RMSEA Parsimony Correction Index

Model	RMSEA	90% Confidence Interval (CI)
Bodfish (2000) six-factor model[a]	0.048	0.043 – 0.053
Lam & Aman (2007) five- factor solution[b]	0.060	0.055 – 0.065
Mirenda et al. (2010) five-factor model[a]	0.092	0.088 – 0.097
Bishop et al. (2013) five-factor model[b]	0.121	0.116 – 0.125
Martínez-González & Piqueras (2017) six-factor model[a]	0.047	0.042 – 0.052
Russell et al. (2019) four- factor model[b]	0.132	0.128 – 0.136
Sturm et al. (2022) six-factor model[a]	0.048	0.043 – 0.053
Study one EFA five-factor model[a]	**0.047**	**0.042 – 0.052**

Note. Only five of the listed models used the full 43-item set from the original RBS-R. These models are depicted below in superscript.

[a]Indicates models using the full 43 items from the original RBS-R.

[b]Indicates models using less than the full 43 items from the original RBS-R, omitting at least one or more items from the tested model.

Comparative fit indices were also used to examine the fit of the proposed model from study one alongside existing factor models. Results for all models are presented below in Table 22. These indices assess a model's fit when compared to a restricted nested model. Study two included the use of the CFI and TLI. CFI values greater than or close to .95 are considered reasonably well-fitting (Hu & Bentler, 1999), and values ranging between .90 and .949 are considered acceptable (Brown, 2006). The TLI is similar to the CFI, though it includes a penalty for more complex models; it also utilizes the same cutoff values as the CFI listed above. The CFI and TLI values for the hypothesized model from study one suggested an acceptable fit (CFI = 0.927 and TLI = 0.922), set by Brown (2006). Across all models examined in study two, three other models of the remaining seven fell within the same acceptable range set by Brown (2006). These models included the Bodfish et al. (2000) six-factor model, the Mirenda et al., (2010) five-factor model, the Sturm et al. (2022) model, and the Martínez-González and Piqueras (2017) six-factor model. The other four models obtained CFI and TLI values <.90, and in some instances, considerably below this cutoff value.

Table 22. CFA Model Results: Comparative Fit Indices

Model	CFI	TLI
Bodfish (2000) six-factor model[a]	0.924	0.920
Lam & Aman (2007) five- factor solution[b]	0.897	0.889
Mirenda et al. (2010) five-factor model[a]	0.719	0.702
Bishop et al. (2013) five-factor model[b]	0.579	0.548
Martínez-González & Piqueras (2017) six-factor model[a]	0.927	0.922
Russell et al. (2019) four- factor model[b]	0.416	0.388
Sturm et al. (2022) six-factor model[a]	0.924	0.920
Study one EFA five-factor model[a]	**0.927**	**0.922**

Note. Only five of the listed models used the full 43-item set from the original RBS-R. These models are depicted below in superscript.

[a]Indicates models using the full 43 items from the original RBS-R.

[b]Indicates models using less than the full 43 items from the original RBS-R, omitting at least one or more items from the tested model.

Summary. To address the question of a reasonable model fit for the hypothesized factor solution, multiple indices were chosen to thoroughly depict how the study one, five-factor model performed across analyses. The model was assessed across several different types of fit indices for this study (i.e., absolute fit, parsimony correction, and comparative fit). According to the major parsimony corrected index (i.e., RMSEA) and the two comparative fit indices (i.e., CFI and TLI), the study one EFA model adequately fit in the CFA validation sample. Not unexpectedly, the absolute fit indices were less supportive (i.e., the chi-square test was significant and SRMR was above .08). For both indices, some absolute deviation was anticipated without cross-loadings included in the CFA model and the high sensitivity of the chi-square test to minor, non-substantive deviations from absolute fit are well known (Bull et al., 1992; Foldnes & Olsson, 2015; Steiger et al., 1990). Overall, it was concluded that, despite non-substantive deviations from absolute fit, the five-factor solution adequately fit the RBS-R variance-covariance matrix of the second sample. Therefore, based on the RMSEA, CFI, TLI, and chi-square /*df* findings, hypothesis 6 was supported.

Research Question 7

In the confirmatory sample, where special education staff rate students with ASD, how does the factor solution generated by the EFA of the RBS-R compare in terms of relative fit to the previous RBS-R factor models found in ASD samples (e.g., Bodfish et al., 2000); Martínez-González & Piqueras (2017); Mirenda et al., 2010; Sturm et al., 2022)?

Hypothesis 7. *It is predicted that the most interpretable model generated by the EFA in sample one will demonstrate a better relative fit to the inter-item covariance matrix of the CFA sample when compared to previously reported factor models derived from ASD samples.*

To complete further comparison of the several models' levels of fit amongst each other, the AIC and BIC indices were analyzed to make a direct comparison between the non-nested models on the same set of data. Because the WLSMV estimator in Mplus does not provide information criterion fit indices (i.e., AIC and BIC), the model was run using the MLR to provide this additional information to assess model fit when compared to any competing models. Only models using the same data set (e.g., the same number of items included in the factor model) were included in the AIC and BIC calculation. Therefore, additional models (Bishop et al., 2013; Lam & Aman, 2007; Russell et al., 2019), based on fewer items, were excluded from the comparison.

However, because of the popularity of the Lam and Aman (2007) factor model, it was considered important to find a way to include it in some way for cross-model comparisons. However, the fact that items were deleted from the original set of 43 items, in order to refine the model, worked against this. As a provisional solution to this model comparison problem, the pattern matrix included in Lam's (2004) dissertation was used to add the missing items back into the Lam and Aman (2007) five-factor model. This allowed the model to have some basis for comparison with other models based on the complete set of 43 items. (In the published model, items 21, 23-35, and 29 were excluded from the factor model due to low loadings and/or issues with cross-loadings. In the reconstructed model, all items were assigned to the factor on which they loaded the highest according to the original 43-item pattern matrix. that they loaded the highest on. Therefore, items 21, 23, 24, and 25 were assigned to Factor I, and item 29 was assigned to Factor IV.) In general, the lower the value of the AIC and BIC, the better the fit of the model. With the reconstructed model, generated for the purposes of comparison, results indicated AIC = 20125.393 and BIC = 20163.933.

The five-factor model retained in the study one EFA resulted in the lowest value for both the AIC and BIC (i.e., AIC = 20213.511, BIC = 20152.051) compared to all other models (see Table 23 for AIC and BIC values for each model). . Following this model, the Sturm et al. (2022) five-factor model was the second highest, with results of AIC = 20249.637 and BIC = 20288.177, followed closely by the Mirenda et al. (2010) five-factor model. Raftery (1995) suggested that differences in BIC values greater than 10 points are considered to be "very strong," or substantial differences. Given the 36-point difference between BIC values of the study one EFA five-factor model and the next-lowest BIC (Sturm et al. (2022) five-factor model), there is clear evidence that, in the validation sample, the retained EFA model from study one fit substantively better than the other models available for comparison. Therefore, hypothesis 7 was supported.

Table 23. CFA Model Results: AIC and BIC Primary Correction Indices

Model	AIC	BIC
Bodfish et al. (2000) six-factor model	20424.145	20464.072
Mirenda et al. (2010) five-factor model	20260.090	20299.534
Martínez-González and Piqueras (2017) six-factor model	20346.950	20386.045
Sturm et al. (2022) five-factor model	20249.637	20288.177
Study one EFA five-factor model	**20213.511**	**20252.051**

Additional Analyses. In addition to the fit indices generated in the CFA, WLSMV parameter estimates, standard errors, two-tailed *p*-values, R^2 values, and residual variances were calculated and are included in Table 24 for the retained and cross-validated five-factor model. (Other examined factor models, including the Russell et al. [2019] four-factor model, the five-factor models from Lam and Aman [2007], Bishop et al. [2013], Mirenda et al. [2010], and Sturm et al. [2022], and six-factor models from Bodfish et al. [2000] and Martínez-González and Piqueras [2017] are included in Appendix D.) Path diagrams for each of the five factors of the five-factor model were generated. Each path diagram includes item loadings and error variances (see Figures 3 through 7). For simplicity and clarity, each factor with its corresponding items was included on its own page, excluding correlations between factors (see Figure 8). Inter-factor correlations generated in the CFA are included separately in Table 25.

Table 24. Study Two CFA Five-Factor Model Parameter Estimates, Standard Errors, R^2, and Residual Variance

Factor	Item #	Item Stem	Parameter Estimate	Standard Error (S.E.)	Parameter Estimate/ Standard Error	R^2	Residual Variance
Insistence on Sameness	39	Insists that specific things take place at specific times	0.900	0.025	36.388	0.810	0.190
	38	Insists on same routine, household, school, or work schedule everyday	0.864	0.026	33.220	0.746	0.254
	33	Insists on sitting at the same place	0.679	0.055	12.238	0.461	0.539
	34	Dislikes changes in appearance or behavior of the people around him/her	0.652	0.068	9.593	0.425	0.575
	37	Resists changing activities; Difficulty with transitions	0.659	0.050	13.211	0.434	0.566
	30	Objects to visiting new places	0.617	0.057	10.766	0.381	0.619
	28	Communication/social interactions	0.625	0.050	12.470	0.391	0.609
	32	Insists on walking in a particular pattern	0.782	0.051	15.209	0.612	0.388
	26	Travel/Transportation	0.904	0.028	32.886	0.818	0.182
	35	Insists on using a particular door	0.886	0.052	16.886	0.784	0.216
	41	Strongly attached to one specific object	0.644	0.049	13.018	0.414	0.586
	27	Play/leisure	0.808	0.032	25.032	0.652	0.348
	40	Fascination, preoccupation with one subject or activity	0.673	0.041	16.248	0.453	0.547
	31	Becomes upset if interrupted in what he/she is doing	0.622	0.050	12.447	0.387	0.613
	20	Hoarding/saving	0.691	0.058	11.994	0.478	0.522
	36	Likes the same CD, tape, record, or piece of music played continually; Likes same move/video or part of movie/video	0.678	0.054	12.561	0.459	0.541
Comp-ulsive Behavio	17	Washing/cleaning	0.625	0.076	8.238	0.391	0.609
	16	Completeness	0.843	0.034	24.621	0.711	0.289
	15	Arranging/ordering	0.831	0.030	27.849	0.691	0.309

Table 24 (cont'd)

	18	Checking	0.795	0.068	11.677	0.632	0.368
	19	Counting	0.831	0.065	12.818	0.690	0.310
	21	Repeating	0.797	0.048	16.741	0.636	0.364
	25	Self-Care – Bathroom and Dressing	0.933	0.026	35.833	0.870	0.130
	29	Insists that things remain in the same place	0.821	0.041	19.919	0.674	0.326
	23	Eating/mealtime	0.750	0.038	19.511	0.563	0.437
	24	Sleeping/bedtime	0.801	0.040	19.825	0.642	0.358
Stereotyped Behavior	2	Head rolling & movement	0.571	0.064	8.909	0.326	0.674
	1	Whole body movement	0.558	0.061	9.171	0.312	0.688
	4	Locomotion	0.818	0.036	22.484	0.669	0.331
	5	Object usage	0.774	0.048	16.107	0.599	0.401
	3	Hand/finger movements	0.714	0.043	16.537	0.509	0.491
	6	Sensory	0.830	0.039	21.417	0.689	0.311
	43	Fascination, preoccupation with movement/things that move	0.788	0.046	17.227	0.621	0.379
	42	Preoccupation with part(s) of object rather than the whole object	0.794	0.047	17.038	0.630	0.370
	22	Touch/tap	0.721	0.064	11.294	0.520	0.480
Body-Focused Repetitive	13	Inserts finger or object	0.512	0.122	4.197	0.263	0.737
	14	Skin picking	0.660	0.078	8.482	0.436	0.564
	11	Pulls	0.827	0.071	11.628	0.684	0.316
	12	Rubs or scratches self	0.860	0.065	13.252	0.739	0.261
Self-Injurious Behaviors	9	Hits self with object	0.901	0.021	41.503	0.811	0.189
	8	Hits self against surface or object	0.493	0.055	8.947	0.843	0.157
	7	Hits self with body part	0.488	0.071	6.901	0.701	0.299
	10	Bites self	0.412	0.066	6.256	0.335	0.665

For the retained five-factor model, in the context of the validation sample, the majority inter-factor correlations, based on the WLSMV estimator, were $\geq$ .30. There were three exceptions: Factor I's (Insistence on Sameness) correlations with Factor IV (Body-Focused Repetitive Behaviors) and Factor V (Self-Injurious Behaviors), as well as the correlation between Factor II (Compulsive Behaviors) and Factor V (Self-Injurious Behaviors). Among the substantive correlations, two were in the large or high range (i.e., between $\geq$ .50 and $\leq$ .70), which related Factor I (Insistence on Sameness) with Factor II (Compulsive Behaviors) and Factor IV (Body-Focused Repetitive Behaviors) with Factor V (Self-Injurious Behaviors). All remaining substantive correlations fell between .330 and .409 (i.e., moderate or medium range).

Table 25. CFA Inter-Factor Correlation Matrix for Five-Factor Solution

		Factor I	II	III	IV	V
Factor	I: Insistence on Sameness	1.000				
	II: Compulsive Behaviors	**0.615**	1.000			
	III: Stereotyped Behaviors	**0.330**	**0.402**	1.000		
	IV: Body-Focused Repetitive Behaviors	0.294	**0.358**	**0.379**	1.000	
	V: Self-Injurious Behaviors	0.257	0.219	**0.409**	**0.549**	1.000

Figure 2. Path Diagram for Study One Five-Factor Model Insistence on Sameness Factor

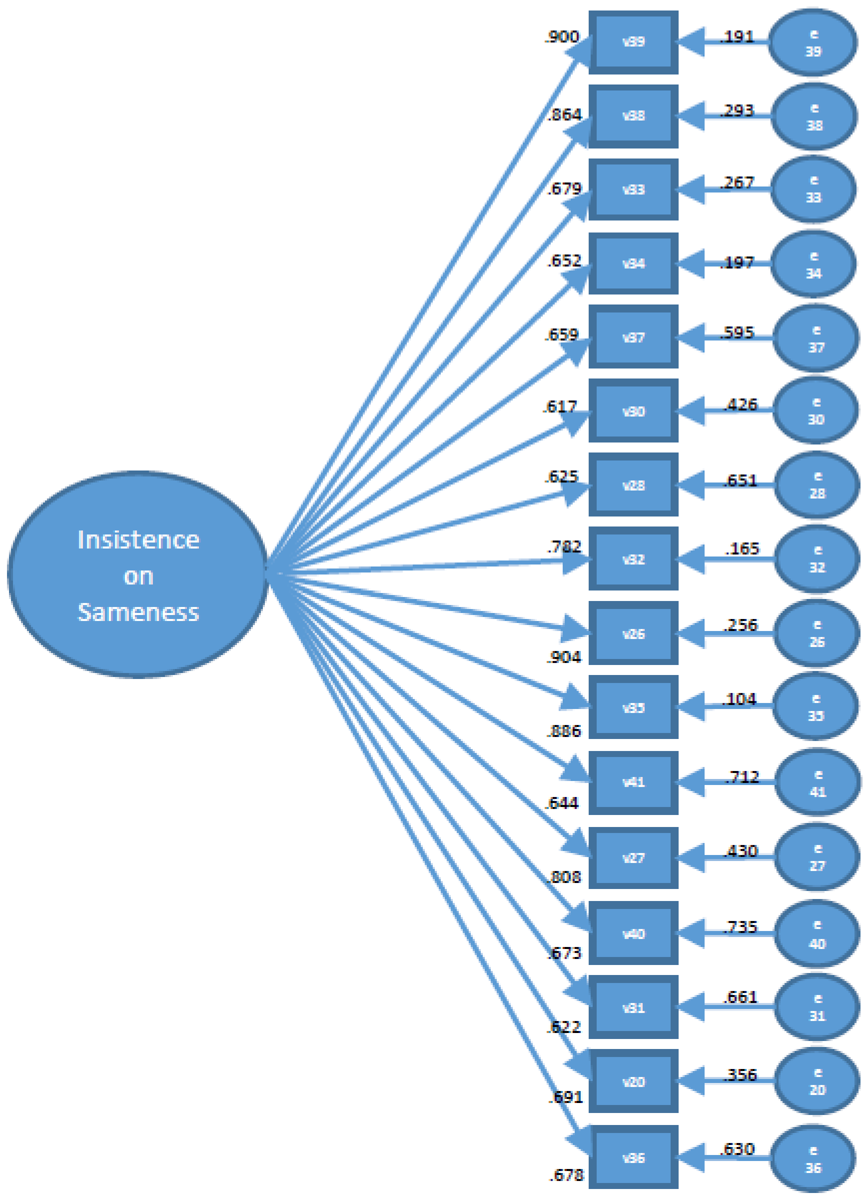


Figure 3. Path Diagram for Study One Five-Factor Model Compulsive Behaviors Factor

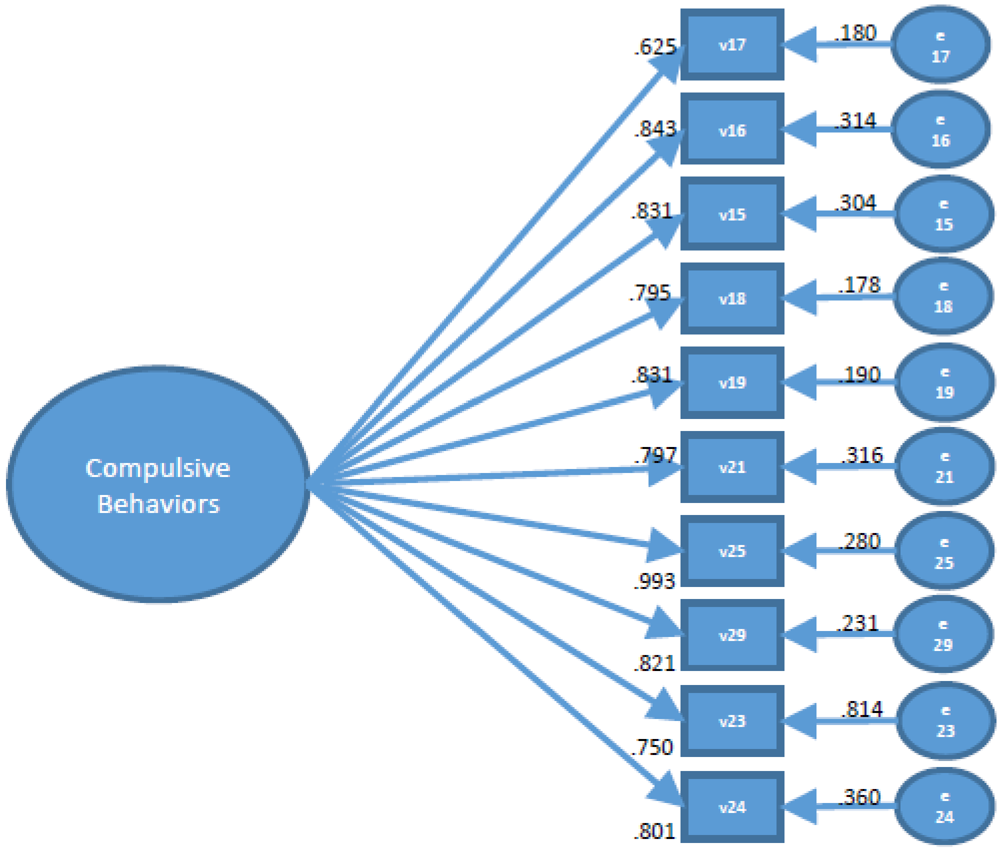


Figure 4. Path Diagram for Study One Five-Factor Model Stereotyped Behavior Factor

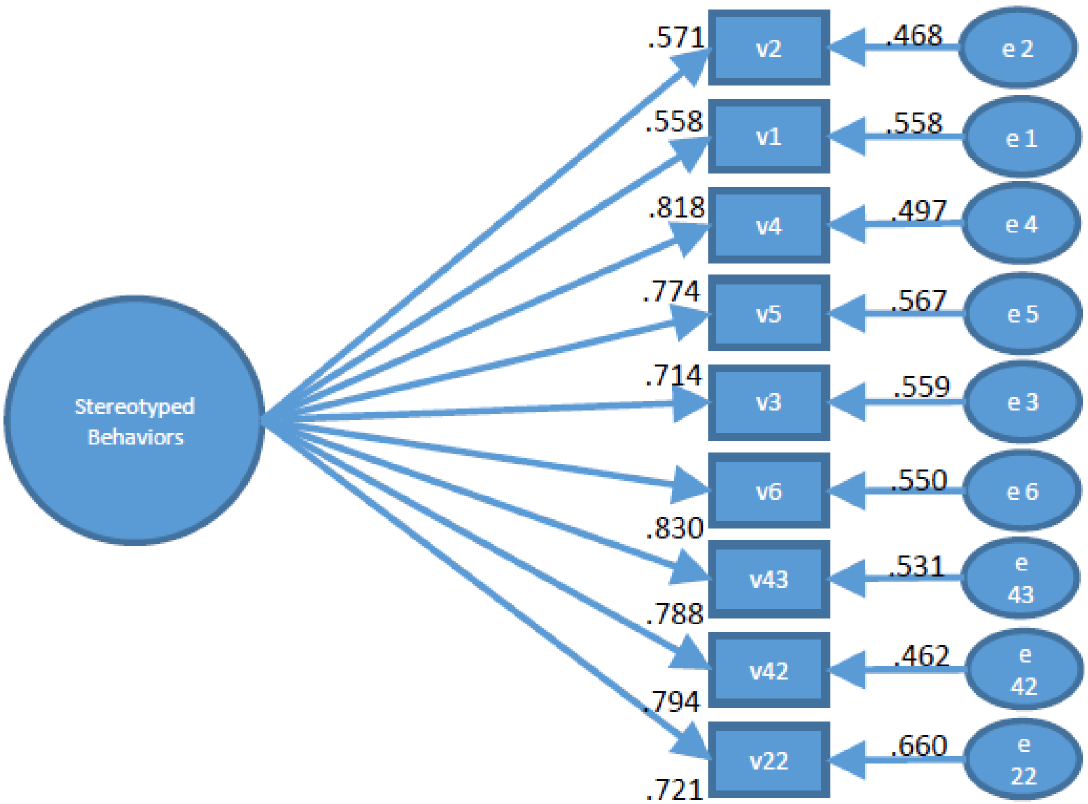


Figure 5. Path Diagram for Study One Five-Factor Model Body-Focused Repetitive Behaviors

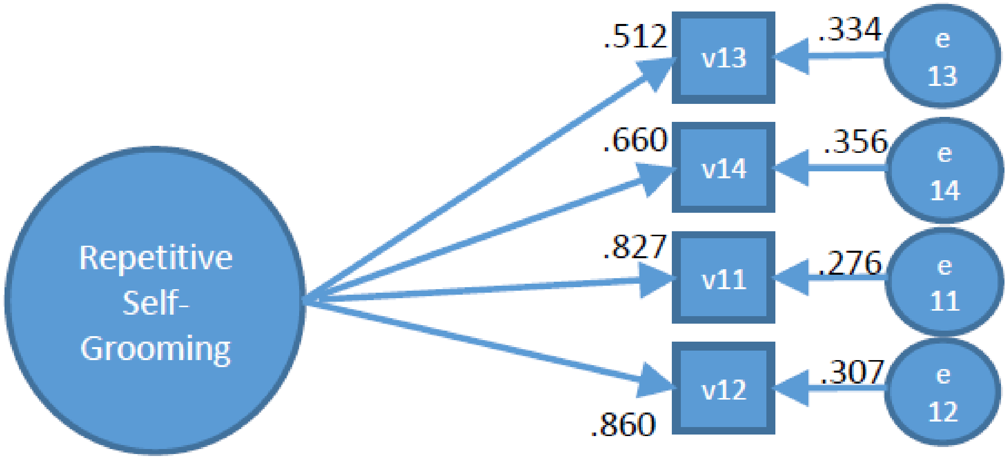


Figure 6. Path Diagram for Study One Five-Factor Model Self-Injurious Behaviors Factor

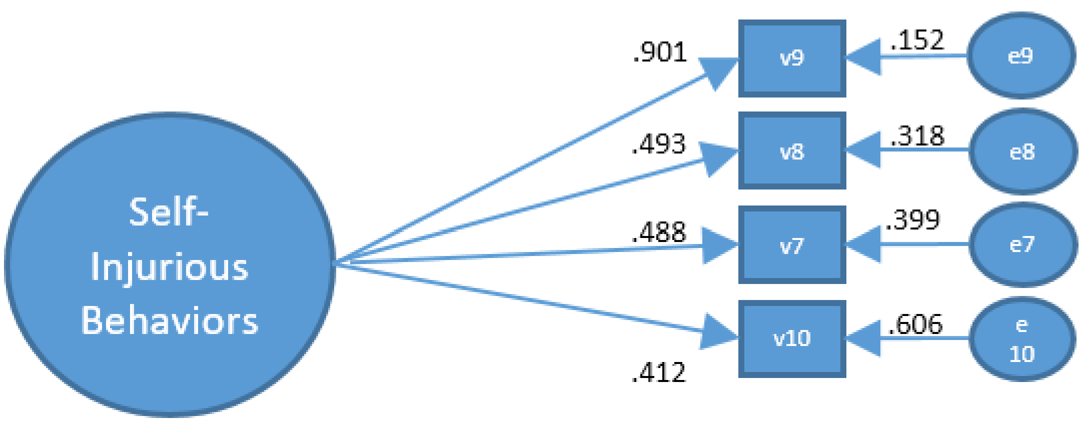
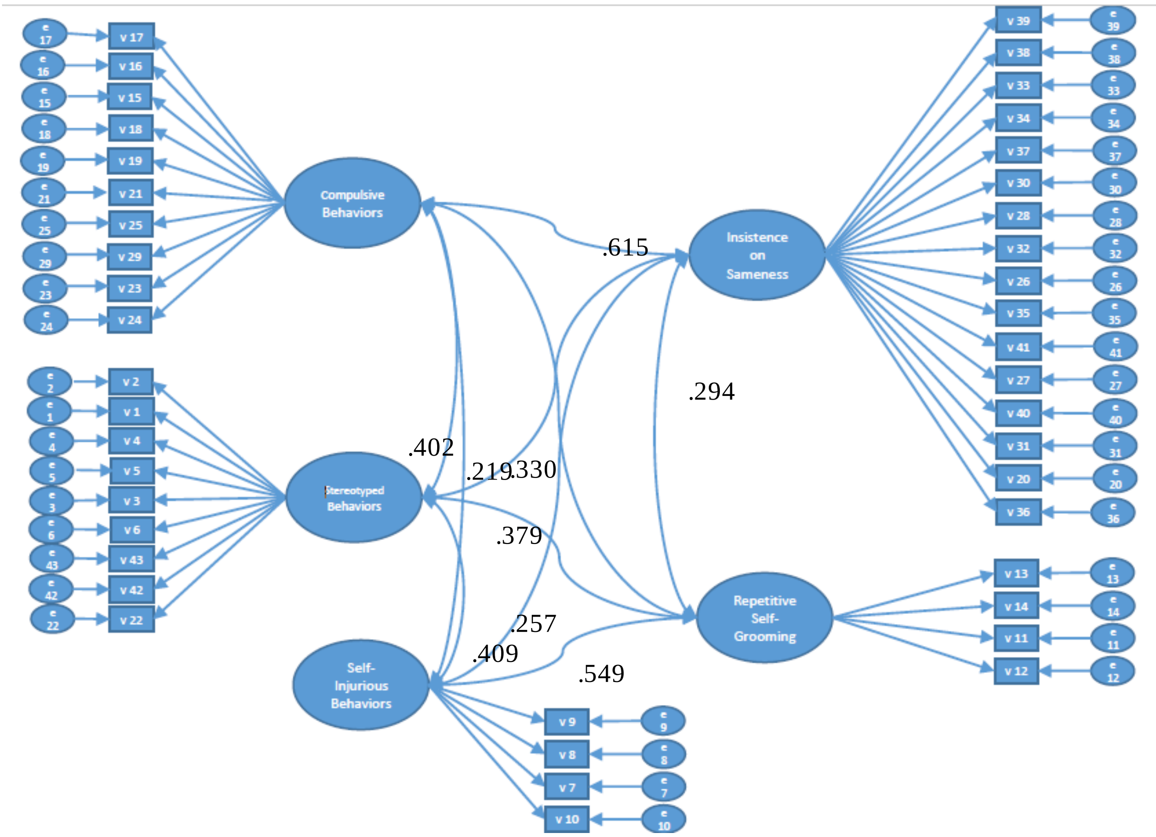


Figure 7. Path Diagram for Study One Five-Factor Model



Summary of Results

All research hypotheses and questions were addressed using two samples of individuals with ASD, rated by teacher and special education staff, across two studies. The tables below (see Table 26 through Table 27) provide a summary of each research question, methods and analyses used, and results from each. Further explanation of each finding will be provided in the discussion (chapter five) of this dissertation.

Table 26. Study One Results Summary Table

Research Question/ *Hypothesis*	Analysis/*Method*	Results
1. When special education staff rate children with ASD, how many possible or likely interpretable factors are available for retention consideration?	EFA with PAF *Scree plot, Kaiser criterion, Velicer's MAP, parallel analysis (PA)*	Velicer's MAP and PA (the most reliable criteria for factor retention) suggested retention ranging between 4 to 6 factors, while scree and Kaiser criterion suggested 3 to 8 factors.
2. When special education staff rate children with ASD, how many factors should be retained to derive the most interpretable factor solution? ***2a.*** *The most interpretable EFA solution will consist of multiple factors (i.e., $\geq$ 2 factors).* ***2b.*** *Within the most interpretable EFA solution, a Self-Injurious Behavior factor will be present.* ***2c.*** *Within the most interpretable EFA solution, a Compulsive Behavior factor will be present.*	EFA interpretive procedure *Four researchers independently interpreted the potential solutions and retained the most interpretable factor solution by consensus.*	The five-factor model was considered the most interpretable solution, unanimously. The five-factor model included both a Self-injurious Behaviors factor and a Compulsive Behaviors factor. (**Hypothesis 2a, 2b, and 2c were all supported**).
3. When special education staff rate children with ASD, does the most interpretable factor structure yield substantive correlations between at least some of the factors of the RBS-R?	EFA with oblique rotation *Examination of the inter-factor correlation matrix for determination of the presence of substantive correlations*	Four of five factors were substantively correlated with each other, while one factor (Body-Focused Repetitive Behaviors) had non-substantive correlations with other factors.
4. When special education staff rate children with ASD, to what extend does the six-factor RBS-R solution from the present EFA (regardless of whether this was the retained solution) correspond to the six-	Qualitative comparison, calculation of the percentage of overlapping items per factor	The five-factors were generally consistent with those identified by Bodfish et al. (2000) for four of six factors. Overall, 25 items (out of

Table 26 (cont'd)

subscale structure originally proposed by Bodfish et al. (2000)? 5. When special education staff rate children with ASD, to what extent does the five-factor RBS-R solution from the present EFA (regardless of whether it was the retained solution) correspond to the five-factor solution reported by Lam and Aman (2007)?	*Examination of the factor constructs of the six-factor solution, compared to the Bodfish et al. (2000) six-factor solution and Lam and Aman (2007) five-factor solution.*	43 total items) overlapped on factors between the two models. The four of the five factors were highly consistent with four of the five identified by Lam and Aman (2007) for four of the five factors. Overall, 30 items overlapped on factors between the two models. Note – Lam and Aman (2007)'s model only included 40 items total, while the EFA-generated five-factor model used all 43 items.

Table 27. Study Two Results Summary Table

Research Question/ *Hypothesis*	Analysis/*Method*	Results
6. When special education staff rate children with ASD, does the mode; generated by the EFA of the RBS-R in the exploratory sample fit the inter-item covariance matrix of the CFA validation sample?? ***6.** The study one EFA solution will reasonably fit the confirmatory sample inter-item covariance matrix.*	CFA with WLSMV *χ2, SRMR, RMSEA, CFI, TLI*	The retained factor solution demonstrated a good fit with the confirmatory sample, for SRMR, RMSEA, CFI, and TLI criterion estimates. Therefore hypothesis 6 was supported.
7. In the confirmatory sample, where special education staff rate children with ASD, how does the factor solution generated by the EFA of the RBS-R compare in terms of	CFA with WLSMV, supplemented by MLR *AIC and BIC (from MLR)*	The retained factor model could not be directly compared using the Mplus DIFFTEST because models differed in the placement of items across models.

Table 27 (cont’d)

absolute and relative fit to the previous RBS-R factor models found in ASD samples (e.g., Hooker at al., 2019; Martinez-Gonzalez & Piqueras, 2017; Mirenda et al., 2010)? ***7.*** *The study one EFA model will demonstrate a better relative fit to the inter-item covariance matrix of the CFA sample when compared to previously reported RBS-R factor models derived from ASD samples.*		The AIC and BIC values for the retained EFA model were lower than all other models in the CFAs, suggesting a better fit for the EFA-generated five-factor model from study one. Therefore, hypothesis 7 was supported.

CHAPTER FIVE: DISCUSSION

Overview of Study One and Study Two

Broadly, study one and study two sought to examine the factor structure of the Repetitive Behavior Scale–Revised (RBS-R; Bodfish et al., 2000) within a sample of individuals with autism spectrum disorder (ASD), as rated by special education staff members. Although several studies have examined the validity and reliability of the RBS-R, no studies used the original (English language) version of the instrument, and no studies have explored its factor structure using ratings from any population beyond caregiver/parent reports. The RBS-R is used frequently as an outcome or research qualification measure, completed by caregivers (Hooker et al., 2019; Mirenda et al., 2010; Scahill et al., 2015; Schertz et al., 2016). However, researchers have highlighted the potential use of this rating scale for other purposes. Some researchers have suggested the use of the RBS-R as a metric of restricted and repetitive behavior (RRB) symptom severity (Travers et al., 2013; Mehling & Tassé, 2016; Martínez-González & Piqueras, 2017). Further, the use of the RBS-R in a school or treatment-based (e.g., Applied Behavior Analysis; ABA) setting may provide useful information about specified treatment goals targeted at reducing core features of ASD, such as RRBI-focused early interventions (Restivo et al., 2005; Wolff et al., 2016).

Though the reliability and validity of the RBS-R have been examined across several studies, validity evidence supporting the most appropriate factor structure of the instrument is conflicting. Overall, the validity evidence supports the use of the RBS-R for aiding in the determination of the severity of RRBs as part of an ASD diagnosis (Mehling & Tassé, 2016). However, researchers have historically suggested the need for a more unified consensus regarding the factor structure of the RBS-R (Mehling & Tassé, 2016; Martínez-González &

Piqueras, 2017; Scahill et al., 2015; Sturm et al., 2022). Since Lam (2004) and subsequently Lam and Aman (2007) published the first EFA of the RBS-R, several studies followed with further EFAs and CFAs of the instrument. First, Mirenda et al. (2010) conducted a CFA of the RBS-R utilizing various theoretical models that were present in the literature at the time, publishing a five-factor model that was similar to that of Lam and Aman (2007)'s five-factor model, but different in the number of items assigned to particular factors. Bishop et al. (2013) published an EFA and CFA of the instrument, settling on a five-factor solution that was like that of Lam and Aman (2007), except with some variance of several items in their own solutions and models. Hooker et al. (2019) published a CFA similar to that of Mirenda et al. (2010) but used more intricate and updated analyses in their CFA – also finding support for (Mirenda et al. [2010]'s) five-factor solution. Russell et al. (2019) also conducted an EFA of the RBS-R, finding support for a four-factor solution of the instrument in. a sample of caregiver raters. This factor solution was like past five-factor models but dissolved the "restricted behavior" factor amongst other factors in the solution. Finally, Sturm et al. (2022) conducted a CFA of the RBS-R, suggesting a good fit for their five-factor model generated from the study's EFA. Across all CFA studies (Hooker et al., 2019; Mirenda et al., 2010; Sturm et al, 2022), a range of one- to six-factor models were examined for fit, though some version of a five-factor model was ultimately landed upon for each study.

All studies listed above utilized the English version of the RBS-R used a population of individuals with ASD, as rated by parents and caregivers. However, only one study (Martínez-González & Piqueras, 2017) examined the instrument using ratings of individuals with ASD made by teachers and education staff. Through this study, the authors found evidence to support the use of a six-factor solution of the RBS-R. However, this six-factor solution differed from the originally published RBS-R, as items were assigned to factors in a slightly different manner. Although this study was the first of its kind to use teacher ratings of participants, it leaves some questions about the factor structure of the RBS-R with such raters, as the study utilized a non-English version of the instrument. To date, there have been no studies utilizing either teachers or special education staff as raters with the originally published (English) version of the RBS-R using a sample of children with ASD.

Thus, this dissertation sought to launch the addressing of the gap in the literature aimed at using raters beyond the parent and caregiver population. Together, the two studies included in this dissertation contributed to the existing literature in terms of independent research on the RBS-R, with some findings supporting aspects of the instrument's validity with teacher ratings. Study one included an EFA on the RBS-R with an ASD sample with special education staff raters. These processes involved the use of thorough EFA methods, including the use of the best practices and theoretical understanding of EFA to guide factor retention. Factor interpretability, potential clinical usefulness in the ASD population, and a general understanding of ASD diagnostic criteria also guided settling on the most appropriate factor structure. Study two involved the use of a CFA on the RBS-R to determine the absolute and relative fit of the study one EFA model. This study also included comparisons of the study one EFA model to existing RBS-R models in the literature using samples of individuals with ASD. The CFA in study two included an extensive comparison of all models in the literature to date and allowed for a direct comparison of non-nested models.

Both studies one and two were intended to clear some of the discrepancies regarding the most appropriate factor structure of the RBS-R. Further, these studies sought to fill in the gap in the existing literature regarding the most appropriate factor structure of this instrument using teacher ratings for individuals with ASD. These findings have at least tentative implications for practice, and when taken together with overall study limitations, provide directions for future research to make clearer directions for use in practical settings. The discussion of the results of studies one and two are provided below – including both summaries and interpretations of findings through the lens of strengths, limitations, and implications for future use and research purposes.

Summary and Interpretations of Findings for Study One

Study one examined the RBS-R using special education staff raters in a sample of 234 students with ASD. This study utilized an EFA with principal axis factoring (PAF) as well as oblique rotation. Multiple criteria were used to best understand the most appropriate number of factors to retain, including the use of a scree plot, Kaiser criterion, Velicer's MAP, and parallel analysis. Research questions one through four are included for discussion below.

Research Question one focused on the number of potential interpretable RBS-R factors for retention using the criteria listed above. With this, solutions ranged from three to eight factors. However, it is noted that the more reliable criterion (e.g., Velicer's MAP and parallel analysis) suggested four to six factors, which was considered most heavily given their reputation for greater accuracy. To be thorough, a range of two to seven factors (plus or minus one factor from the more reliable standards and the addition of the two-factor model to be consistent with past studies such as Georgiades et al., 2010 using the Greek version of the RBS-R) were included for further examination.

The decision to consider two- to seven-factor solutions aligns with the number of solutions examined in existing studies present in the literature. Across the RBS-R EFA and CFA literature, researchers (Hooker et al., 2019; Lam & Aman, 2007; Mirenda et al., 2010; Russell et al., 2019; Sturm et al., 2022) examined from a range of one- to six-factor solutions for retention. In this study, a one-factor solution was not considered for retention, as RRBIs have generally been determined to be multidimensional (Bishop et al., 2013; Lam & Aman, 2007; Mirenda et al., 2010; Sturm et al., 2022; Turner, 1999). More specifically, Lam and Aman (2007) reported that the two- to six-factor solutions were deemed appropriate to consider for retention, meeting their predetermined criteria (see below). Further, past studies such as Russell et al. (2019) selected two- to five-factor solutions for further examination.

This study also utilized best practices for factor retention criteria, which included a combination of methods utilized by existing studies in the literature. Lam and Aman (2007) used the scree plot method, Eigenvalues above 1.0, RMSEA values for Goodness-of-Fit, and clinical interpretability. Bishop et al. (2013) report using chi-square and comparison of RMSEA values. Like Lam and Aman (2007), Russell et al. (2019) report using Eigenvalues above 1.0, scree plot, and clinical interpretability. Research question one sought to select the number of retained factors using scree plot, Kaiser criterion, Velicer's MAP, and parallel analysis. Including the use of more methods allowed for a less restricted range of options. Overall, the use of the listed four retention tests led to more comprehensive and interpretive information, compared to previous EFA studies of the RBS-R – ultimately yielding results suggesting the examination of two to seven factors.

Research question two focused on the interpretability of the derived factor solution of the RBS-R from research question one. Pattern matrices went through an oblique rotation and were then assembled into factor models. Again, factor solutions between two and seven factors were considered, though only the four-, five-, and six-factor solutions appeared most viable based on several factors set by criterion tests used in research question one. A team of four independent researchers considered all factor solutions (two-, three-, four-, five-, six-, and seven-factor solutions). However, most consideration and attention were focused on the four-, five-, and six-factor solutions based on the parallel analysis, MAP test, and additional tests from research question one. Consensus between all four of the researchers was reached that the five-factor solution was the most interpretable and provided the most support for clinical interpretation, beyond the criterion listed above. Hypotheses generated for Research Question Two were that 2a) the most interpretable EFA solution will consist of multiple factors (i.e., $\geq$ two factors); 2b) a Self-Injurious Behaviors factor would emerge, and 2c) a Compulsive Behaviors factor would emerge. All hypotheses for research question two – hypotheses 2a, 2b, and 2c, were supported.

Ultimately, the five-factor solution was settled on after consideration of study one data, as well as theory relating to factor analyses and RRBIs. When considering the clinical meaningfulness of the factor solution, researchers considered the constructs themselves – if they were clearly defined and consistent with the pillars of RRBIs. The solution fell within the range of factors supported by tests included in research question one and was the solution with the fewest number of cross-loadings for items to factors. The five-factor solution included factors with an appropriate number of items on each factor, unlike the six- and seven-factor solutions which included factors with two or three items within the structure.

The four-factor solution, which was also examined closely by the team of independent researchers, has some support for retention based on the tests completed under research question one. However, additional issues hindered this solution from further consideration. One issue that led to the eventual elimination of this solution was due to its high number of cross-loadings. Further, this factor solution included a few items (e.g., items 36, 22, and 23) with no unique loading on a factor. One of the most substantial issues with the four-factor solution is the first factor (Factor 1), which combines the domains of sameness and ritualistic behavior. This factor includes a broad range of items that may not completely capture one unique construct. That is, most items (e.g., items 39 and 38 that load the highest on this factor) represent an insistence on sameness, with loadings of .91 and .90, respectively. This factor includes items that have historically been compulsive in other studies (e.g., items 20, 23, 24, and 29), despite having a factor (Factor II) that represents compulsive behaviors wholly.

Interestingly, one of the most interesting observations that came from the examination of the four-, five- and six-factor solutions was the clear emergence of the Stereotyped Behaviors factor. All factor solutions generated from the EFA in study one except for the two-factor solution, (i.e., the three-, four-, five-, six- and seven-factor solution) produced a clear, identical Stereotyped Behaviors factor with the same factors across solutions. This Stereotyped Behaviors factor includes items 1, 2, 3, 4, 5, 6, 22, 42, and 43, which is also identically present in studies in the current literature (Lam & Aman, 2007; Martínez-González & Piqueras, 2017; Sturm et al., 2022). Additional models (e.g., Bishop et al., 2013; Russell et al., 2019) had the same factor structure with identical items, aside from items that were omitted from the overall models – including item 22 from both studies, as well as item 42 in Bishop et al. (2013)'s model. The clear presence of this factor across not only the study one EFA solutions, but also the published models in the literature, suggest that stereotyped behaviors are represented well on the RBS-R.

Perhaps this strong representation stems from the inclusion of the term "stereotyped" in the DSM-5 (APA, 2013) definition of ASD, while all other categories of repetitive behavior on the RBS-R are more discreetly described in the DSM-5. The inclusion of the term "stereotyped" and less- apparent inclusion of other RBS-R constructs is consistent in the DSM-IV-TR (APA, 2000) and DSM-IV (APA, 1994) – both of which, including the DSM-5, were used across the studies that have yielded a strong Stereotyped Behavior factor. Though not explicitly used in these studies, the ICD-10 criteria for autism include the same pattern of explicitly stating the "stereotyped behaviors," while leaving other constructs of RRBIs within the RBS-R less clear. Stereotyped behaviors are also very present in other measures of RRBIs. A clear Stereotyped Behaviors factor in some forms is present on the Behavior Problems Inventory (BPI-01: Rojahn et al., 2000), the Pervasive Developmental Disorder Behavior Inventory (PDDBI: Cohen et al., 2003), and Repetitive Behaviours Questionnaire 2 (RBQ2: Leekham et al., 2007).

Aside from the strength of the Stereotyped Behaviors factor on the four- and six-factor solutions, the team of independent researchers settled upon the five-factor solution. Though this was not hypothesized to be the solution retained (as hypothesis 2c stated that at least a two-factor solution would be retained), it was not surprising that this solution was selected. As there has been ample support for a five-factor solution across the literature, the five-factor solution from the study one EFA boasts a few notable differences from the existing five-factor models. Namely, the emergence of the Body-Focused Repetitive Behaviors factor stands out in the selected solution, as well as the dissolution of the items formally attached to "restricted behaviors" broadly. The explanation of the changes present may in part be because this is the first study using the original English version of the RBS-R with a sample of teacher raters, as opposed to previous studies using either different raters or a foreign-language version of the instrument. This leads to the question of whether this newly proposed five-factor solution is a better fit for the sample and sample raters and how it can be generalized in future studies. Analysis of the inter-factor correlations aids in further exploration of the factor constructs present in this solution, though there is strong evidence to support the presence of this emerging factor. Further, estimates and interpretation of this model's absolute and relative fit are discussed in study two of this dissertation (below).

Research question three focused on the strength of the inter-factor correlations for the five-factor solution. Four of the five factors were substantively correlated with each other, as determined through the standard set in the factor analysis criteria that correlations of > .30 are considered substantive (Beavers et al., 2013; Nunnally & Bernstein, 1994; Pett et al., 2003). Factor I (Insistence on Sameness) was largely correlated with Factor II (Compulsive Behaviors), Factor III (Stereotyped Behaviors), and Factor V (Self-Injurious Behaviors). Factor III (Stereotyped Behaviors) was also correlated moderately with Factor V (Self-Injurious Behaviors). One factor (Factor IV), Body-Focused Repetitive Behaviors, was the only factor that did not yield any inter-factor correlations that were above .30. This factor had correlations of .10 (Factor I), .16 (Factor II), .22 (Factor III) and .28 (Factor V). While some of the correlations closely approached the .30 criteria, they were not correlated enough to meet this standard.

Internal consistency reliability estimates were also calculated using both ordinal and Cronbach's alpha. Ranges of internal consistency estimates ranged from .871 to .937 for ordinal alpha and .720 to .873 for Cronbach's alpha. Regardless of the reliability estimate considered, the subscales all met the minimum standard for research purposes (set by Nunnally, 1978) when examining both Cronbach and ordinal alphas. Both alpha estimates also suggested that all scales met the standard set by Murphy and Davidshofer (1998) adequate for weekly progress monitoring ($\geq$ .70). As previously stated in Chapter 3, ordinal alpha was used as the primary estimate of internal consistency, as it attempts to adjust for the ordinality of the item scales by replacing the commonly used Pearson correlations with polychoric correlations in the formula for coefficient alpha. For this study, three of the five subscales (Insistence on Sameness, Compulsive Behaviors, and Stereotyped Behaviors) fell in the excellent range (i.e., .90 to .99) and two subscales (Body-Focused Repetitive Behaviors and Self-Injurious Behaviors) fell in the moderately high or good range. These results highlight the strength of the retained model's subscale, particularly when adjusting for the scale's item ordinality.

Although ordinal alpha is considered the most appropriate statistic of internal consistency reliability for data that is ordinal in nature, such as the RBS-R, all existing studies examining the RBS-R reported calculations of Cronbach's alpha. The current study reports both Cronbach's and ordinal alphas for measures of best practice and provides a means of comparison across the existing factor models. Generally, the internal consistency reliability estimates from the study one EFA five-factor solution align with those in published studies. Lam and Aman (2007) reported Cronbach's alpha measures ranging from .78 (Restricted Interests) to .91 (Rituals/Sameness Behavior) for their five-factor model. Martínez-González and Piqueras (2017)'s estimates ranged from .70 (Compulsive Behaviors) to .88 (Sameness Behaviors) for their six-factor model. Mirenda et al. (2010) reported ranges of internal consistency reliability from .71 (Ritualistic Behavior) to .88 (Sameness Behavior) for their five-factor model. Sturm et al. (2022)'s five-factor model yielded internal consistency estimates ranging from .66 ("Factor V") to .88 ("Factor IV"). Finally, Russell et al. (2019) reported that all factors yielded estimates that were $\geq$ .70. The listed estimates all follow a similar pattern to that of the study one EFA five-factor model. Such high estimates of internal consistency reliability provide evidence that the generated five-factor model is a sustainable model for use with teacher raters of individuals with ASD.

Overall, the inter-factor correlations supported the five-factor solution's structure. Correlations present in the five-factor solution ranged from .10 (Body-Focused Repetitive

Behaviors and Insistence on Sameness) to .56 (Compulsive Behaviors and Insistence on Sameness) – a range of low relationship to large (i.e., correlations $\geq$ .50 are deemed as "large" by Cohen, 1988). None of the inter-factor correlations suggested a strong relationship (i.e., $\geq$ .80), which would suggest that factors may be too redundant – measuring the same construct (Brown, 2006). Therefore, the factor structure of this five-factor solution was supported adequately.

One prominent five-factor solution in the RBS-R literature is the Lam and Aman (2007) factor solution. This five-factor solution included inter-factor correlations that ranged from .14 (Restricted Interests and Self-Injurious Behavior) to .55 (Compulsive Behaviors and (Ritualistic/ Sameness). This range was comparable to the study one EFA five-factor solution, which also included a range of low relationships to large correlations within their published study. Further, both five-factor solutions yielded the highest inter-factor correlations between the same factors – the "sameness" behaviors factor and Compulsive Behaviors. Inter-factor correlations in other studies examining the RBS-R did not include correlations as a result of their EFA (i.e., studies only reported correlations following the CFA methodology) and are therefore not included. Though, it is important to note the similarities in the pattern and range of inter-factor correlations between these two five-factor solutions.

The relationships between factors are important pieces of information, as they provide a deeper analysis of the constructs present within the chosen factor solution. Upon further examination of the Body-Focused Repetitive Behaviors factor, this factor had the lowest correlations with all other factors, yielding correlations all below .30. The correlation closest to .30 within this factor is found within the correlation of Body-Focused Repetitive Behaviors and the Self-Injurious Behavior factor (.28). Conceptually, this makes sense, as the two factors have been placed together under one unified factor in all previous studies (Bishop et al., 2013; Bodfish et al., 2000; Hooker et al., 2019; Lam & Aman, 2007; Martínez-González & Piqueras, 2017; Mirenda et al., 2010; Russell et al., 2019; Sturm et al., 2022) up until this present study. There has yet to be a division of the original self-injury factor from the RBS-R in any study, leaving room for further interpretation and future research in this subset of individuals and raters.

Body-Focused Repetitive Behaviors (BFRBs) are an umbrella name for behaviors related to impulse control. These behaviors involve compulsively damaging an individual's physical appearance, sometimes causing injury (Grant & Stein, 2014). Compared to repetitive behaviors overall, BFRBs are newer in the field of psychopathology and are cited as being strongly considered as their own classification in the ICD-11 (Grant & Stein, 2014). Currently, little is known about well-established interventions that target the treatment of BFRBs (Teng et al., 2004; Woods & Houghton, 2015). BFRBs are frequently associated with anxiety disorders and include related disorders such as excoriation disorder (skin picking), cheek biting, inner lip biting, nail picking, nail-biting, compulsive nose picking, hair biting, hair pulling, and tongue biting (Penzel, 2003; Stein et al., 2008; Mansueto et al., 2020), most of which are represented on this emerging factor – given the items that were originally on the scale embodying such behaviors. BFRBs have been positively associated with core features of ADHD, including executive functioning challenges such as planning and organization difficulties (Flessner et al., 2015). Interestingly, because BFRBs are so strongly associated with anxiety-related disorders, several studies (e.g., Grant & Stein, 2014; Houghton et al., 2018; Robert et al., 2014; Siddiqui et al., 2012) that study BFRBs that explicitly exclude children with ASD or intellectual disabilities – making an interesting distinction between this subset of self-injurious behaviors among populations.

Though there is no research concerning the detection of BFRBs among different raters (e.g., clinicians versus teachers versus caregivers, etc.), the split of the original Self-Injurious Behavior into two new SIB and BFRB factors raises questions about the distinction being made between teacher raters in this study. Some studies have demonstrated differences in parent versus teacher ratings for samples of children with autism on measures of adaptive behavior (Lane et al., 2013), social skills (Azad et al., 2016; Murray et al., 2009), and challenging behaviors (Murray et al., 2009; Rapin et al., 1999). Specific to the core features of ASD, Kaat and Lecavalier (2013) found a poor to fair agreement between parent and teacher ratings on autism-related symptoms. These differences may be due to environmental differences, or differences in the way children present behaviorally at home versus school – which has been discussed specifically in ASD literature (Posserud et al., 2006; Ronald et al., 2008). This may also be due to differences in rater experience. Teachers and caregivers may have different comparison samples in their ratings, as teachers typically have experience working with same-aged peers of those they are rating. Further, teachers have more opportunities to observe differences in the presentations of children with autism in different environments (Jepsen et al., 2012; Ryland et al., 2012). In this study, teachers who rated participants were comparing them to other children with autism. This may have contributed to observed differences and nuances of self-injury and repetitive behavior that have not been noticed by caregivers in other studies, leading to a more nuanced view of the behaviors.

Research question four sought to provide a detailed comparison between the Bodfish et al. (2000) six-factor model and the study one EFA-generated six-factor. Though there was no guarantee that a six-factor solution would be produced from the study one EFA, its presence amongst the previous studies led researchers to believe that there would likely be present, regardless of whether the solution was retained or not. A comparison of these studies was carried out through qualitative comparisons of factor names from both solutions. Among both models, factor names emerged relating to Sameness (Sameness Behaviors [Bodfish et al., 2000] and Insistence on Sameness [Study one EFA]), Compulsive Behaviors, Stereotyped Behaviors, Self-Injurious Behaviors, and Restricted Behaviors. However, in one instance, the original Bodfish et al. (2002) six-factor model proposes a Ritualistic Behavior subscale, whereas the EFA-generated six-factor solution proposes a Body-Focused Repetitive Behaviors factor – a division of the original Self-Injurious Behavior factor, as described in Research Question 3.

In general, there was substantial overlap in the highest loading items across most factors for the first several items. However, the order of items according to the highest-ranking factor solution differed slightly in their structure. For example, the Insistence on Sameness/Sameness Behavior factors shared the same five out of six highest-loading items on the same factor (items 39, 38, 37, 33, and 30). Additionally, a high percentage of items from the Bodfish et al. (2000) six-factor solution were also present in the factor solution from study one. Two sets of analyses were completed to calculate the total item overlap percentage – one with the number of items from the study one EFA six-factor solution in the denominator, and one with the number of items from the published Bodfish et al. (2000) in the denominator (with the number of items overlapping in the numerator for both sets). At the factor-specific level, the Compulsive Behavior(s) subscales included the highest percentage of overlap consistently, with six items overlapping out of eight (75%) for both the study one EFA six-factor model, and the published RBS-R model. For the previously discussed factors with the same name and concepts (e.g., Insistence on Sameness/Sameness Behavior, Compulsive Behavior(s), Stereotyped Behavior(s), Self-Injurious Behavior(s), and Restricted Behavior(s) factors), 68% of items overlapped between the factors. While this number is not particularly high, these models were largely the same, with most differences stemming from the presence of a different factor in each solution.

Several similarities stood out between the two six-factor solutions. Aside from the presence of one unique factor in each model, the Ritualistic, Compulsive, Sameness, Stereotyped, and Self-Injurious Behavior factors were clearly existent across the two solutions. Within these factors, the “Sameness” behaviors, Compulsive Behaviors, and Stereotyped Behaviors all had at least 75% overlap of items across the two models. Across all factors, the Stereotyped Behaviors subscale was the most similar between the two models. The EFA-generated solution included all the same items for this factor that Bodfish et al. (2000) did, but also included three additional items – two of which were originally part of Bodfish et al. (2000)’s Restricted Behavior subscale, which does not exist on the EFA-generated solution. The “Sameness” factors had a 73% and 57% overlap of items. The major difference between the two Sameness-related factors was that three items from the Bodfish et al. (2000) solution were not present (items 29, 35, and 36), and six items (items 26, 28, 24, 27, 23, and 41) were present in the EFA-generated solution that were not present in the original model. The items left out of the EFA-generated solution (items 29, 35, and 36) loaded onto different factors – Compulsive Behavior (item 29) and Restricted Behaviors (items 35 and 36). The additional items that were not present in the published RBS-R but did load onto the Insistence on Sameness factor primarily (five out of six items) came from the Ritualistic Behavior Subscale on the published RBS-R, with the final (item 41) coming from the Restricted Behaviors subscale.

The most notable difference between these two six-factor solutions is the obvious presence of one unique factor on each model (i.e., Body-Focused Repetitive Behaviors on the study one EFA six-factor model and Ritualistic Behavior on the Bodfish et al., 2000 model).

Aside from this, it is important to note that, despite the Self-Injurious Behavior(s) exhibiting the same name, they do not have a high percentage of item overlap. With the division of the published RBS-R's Self-Injurious Behavior (items 7 through 14) into the factors Self-Injurious Behaviors (items 7 through 10, and item 20) and Body-Focused Repetitive Behaviors (items 11 through 14) on the EFA-generated solution, several items were assigned to different factors. Further, the originally published RBS-R's Restricted Behavior seemingly dissolved in the EFA-generated six-factor model, with only one item (item 40) overlapping between the two models – which also loaded the highest on both factors.

Despite differences in the representation of Restricted Behaviors and Self-Injurious Behaviors in both solutions, these constructs are still clearly present within the RBS-R itself. The two models differ greatly in their conceptualization of self-injury, causing it to take on a higher-order and lower-order form in the generated EFA study. This pattern is also present in the comparison of the two five-factor solutions, examined through research question five (below). That said, there is an assumed distinction made by teacher and special education staff raters on the RBS-R for this sample, compared to all the other solutions in published studies using parent and caregiver ratings. Overall comparison of the two six-factor solutions indicates the strong presence of many constructs. However, questions remain about the solidity of these constructs, which were arguably made clearer in the five-factor solution generated in the study one EFA.

Similar to research question four, research question five set out to provide a close comparison between the Lam and Aman (2007) five-factor model and the EFA-generated five-factor model from study one. Though it was unknown that the five-factor solution would be selected as the most desirable solution at the outset of the study, it was determined that examination of Lam and Aman (2007)'s solution would provide helpful information, as the presence of a five-factor model is highly cited throughout the RRB literature. Methods used to answer research question five were much like those used in research question four. This was executed through a qualitative comparison of factor names from both solutions, examination of the highest loading items within each factor, and calculation of the percentage of item overlap for similar factors. Amongst the two individual factor solutions, similar factor names were found in Lam and Aman (2007; Ritualistic/Sameness Behaviors, Self-Injurious Behaviors, Stereotypic Behaviors, Compulsive Behaviors, and Restricted Interests) and in the EFA-generated five-factor solution from study one (Insistence on Sameness, Compulsive Behaviors, Stereotyped Behaviors, Self-Injurious Behaviors, and Body-Focused Repetitive Behaviors). The highest loading items were similar across factors in each solution. Although, these items differed in their order across the two different five-factor solutions. Most items from Lam and Aman (2007) were found in similar factors in the EFA-generated five-factor solution. The chief difference between the two solutions was similar to that of the six-factor solution – the presence of the Body-Focused Repetitive Behaviors factor emerged, which was a division of the broader Self-Injurious Behavior factor from the Lam and Aman (2007) solution.

Broadly, the Lam and Aman (2007) five-factor solution varied in its similarity to the EFA-generated five-factor solution from study one. Again, two sets of analyses were completed to calculate the total item overlap percentage – one with the number of items from the study one EFA five-factor solution in the denominator, and one with the number of items from the Lam and Aman (2007) solution in the denominator (with the number of items overlapping in the numerator for both sets). For the Sameness factors, 12 items overlapped for the study one EFA (75% overlap) and for the Lam and Aman (2007) solution (86% overlap). Study one's Insistence on Sameness factor included three additional items (items 20, 40, and 41) that were not included in Lam and Aman (2007)'s. The first two items on each factor are the same (items 39 and 38), though the rank order starts to differ starting at the third item. Similar results were found in the comparison of the two Compulsive Behavior(s) factors present in both solutions. These factors had an 83% item overlap and 50% overlap, with the same first five items on each factor loading highest on their respective factor. The Stereotyped Behaviors factors were also similar for both solutions. All six items on the Lam and Aman (2007) Stereotyped Behaviors factor were also present on the corresponding factor from the study one solution, with a 100% item overlap for all nine items, with no added or omitted items. The only difference is that the two solutions are in the rank order of the highest five items. That is, the top five highest loadings include the same items, just in a different rank order. The similarity between these three factors (Sameness, Compulsive, and Stereotyped Behaviors) for both studies demonstrates the strong representation of a construct related to sameness, compulsive, and stereotyped behaviors on the RBS-R. This was similar to that found in the comparison of the two six-factor solutions in research question four; however, there is a stronger similarity between the three factors for the five-factor solutions due to higher item overlap and factor-item construction.

Again, the introduction of the Body-Focused Repetitive Behaviors factor in the study one EFA-generated solution led to prominent differences present between the two five-factor solutions. The original (Bodfish et al., 2000) solution's Self-Injurious Behavior included items 7 to 14, which was replicated by Lam and Aman (2007), as they included an identical factor with the same name. However, the EFA-generated five-factor solution from study one only included items 7 through 10. The Body-Focused Repetitive Behaviors factor included the other half of Lam and Aman's Self-Injurious Behavior factor, including items 11 through 14. It is also important to note that Lam and Aman (2007)'s Restricted Interest factor was dissolved in the present study, with all items (36, 40, and 41) loading onto the study one solution's Insistence on Sameness Factor. These items all yielded substantive loadings on the Insistence on Sameness factor. This raises questions about the overall stability of the representation of restricted interests on the RBS-R as this factor was dissolved in the present study using special education staff raters. Similarly, the Restricted Interests factor is consistently the weakest across most solutions in various studies (Bodfish et al., 2000; Lam & Aman, 2007; Martínez-González & Piqueras, 2017; Mirenda et al., 2010) that utilize caregiver ratings. Regardless of the depiction of restricted interests on the RBS-R, findings from research questions four and five suggest that the use of teacher and special education staff ratings may require a different factor solution than those that are supported currently in the literature for parent and caregiver raters.

Summary and Interpretations of Findings for Study Two

Study two examined the RBS-R using special education staff raters in a sample of 233 students with ASD. This study utilized a CFA to examine model fit through analysis and examination of χ^2, SRMR, RMSEA, CFI, and TLI from the Mplus Weighted Least Squares Mean and Variance Adjusted (WLMSV) estimator. Additionally, AIC and BIC indices were calculated for model comparison using the Robust Maximum Likelihood (MLR) estimator on Mplus. Such estimates were calculated for the five-factor solution generated from the study one EFA, as well as for existing RBS-R factor models in the current literature.

Research question six set out to focus on the fit of the EFA-generated model with the validation sample through examination of the inter-item covariance matrix. It was hypothesized that the five-factor model from study one would at least adequately fit the RBS-R variance-covariance matrix of the second ASD sample. A CFA was performed using the Mplus WLSMV estimator to calculate relative fit indices for individual models – including χ^2, SRMR, RMSEA, CFI, and TLI. Regarding fit, the CFA yielded results that suggested the five-factor CFA from study one met or approached the cutoff values for all the indices examined. Therefore, hypothesis six was supported, as the five-factor study one EFA model demonstrated adequate fit to the RBS-R variance-covariance matrix of the sample for study two.

The study one EFA five-factor model stood out against the seven others it was compared against – including the original six-factor Bodfish et al. (2000) and Martínez-González and Piqueras (2017) models, as well as the four-factor Russell et al. (2019) model and five-factor models from Lam and Aman (2007), Mirenda et al., (2010), Bishop et al. (2013) and Sturm et al. (2022) studies. Across the models tested for relative fit, all models yielded a statistically significant χ^2, and therefore did not meet the set criterion. However, larger sample sizes can too easily yield statistically significant results over minor discrepancies. Such results are common in the CFA literature, as the chi-square fit index very strictly assesses exact fit and its power to detect model discrepancies is highly dependent on sample size. context, can too easily yield statistically significant results over minor discrepancies. Through consideration of the standard set by Alvi et al. (2020), which considers degrees of freedom into a ratio of chi-square, the EFA-generated five-factor model fell within the category deemed reasonable, in addition to models from Lam and Aman (2007), Mirenda et al. (2010; and Russell et al. (2019). All models except Bodfish et al. (2000) met the criteria for SRMR cutoff. According to the criterion set for RMSEA standards, the study one, five-factor model indicated a close fit, along with three other factor models that fell into this descriptive category (Martínez-González & Piqueras, 2017; Sturm et al., 2022; and Bodfish et al., 2000). Regarding CFI/TLI standards, the study one EFA five-factor model fell within acceptable limits, in addition to factor models from Bodfish et al. (2000) and Martínez-González and Piqueras (2017), and Sturm et al. (2022). Across all measures of fit, the study one EFA generated results that suggested the closest fit to the standard, compared to all other existing factor models included in the study. However, it is noted that other models, namely, the Martínez-González and Piqueras (2017), and Sturm et al. (2022) models, also consistently fell into similar descriptors.

Research question seven aimed to examine how the EFA-generated factor solution compared to existing RBS-R factor models (e.g., Bodfish et al., 2000; Lam and Aman, 2007; Martínez-González & Piqueras, 2017) in terms of relative and absolute fit. It was hypothesized that the five-factor model from the study one EFA would demonstrate a better fit to the second ASD sample's inter-item variance-covariance matrix than the existing factor models from the published literature. For this research question and hypothesis, the Mplus MLR estimator was used to calculate AIC and BIC values for cross-model comparisons. In preparation for addressing this research question, it was assumed that the Mplus DIFFTEST would be used to test the significance of the difference in fit between two nested models. However, there were no existing models that fit the criteria to be considered as "nested."

AIC and BIC fit test results indicated that the study one EFA five-factor model was the best fitting model, compared to the other existing factor models from the current literature. The Sturm et al. (2022) and Mirenda et al. (2010) models both yielded AIC and BIC results that were within 50 points of the EFA-generated five-factor model. Raftery (1995) states that BIC differences of greater than 10 points should be considered as very strong, or substantive, differences. Given this information, hypothesis seven was fully supported. However, after considering the results from the inter-factor correlation outputs from other models which yielded moderate correlation among factors, there is evidence to support that the other listed five-factor models have some credibility as factor structures for this sample as well. It is noted that only the

AIC and BIC fit indices were used in study two (research question seven) to establish a direct comparison of fit between models. These comparisons were based on the unique variance-covariance matrix, used only in study two itself. Although the use of the Mplus DIFFTEST was not applicable in this study, the inclusion of an additional measure of comparison would provide further information about fit comparison as well.

Overall, the study results from study two appear to parallel results from previous CFAs that compared several existing factor models in their studies using caregiver raters (Hooker et al., 2019; Mirenda et al., 2010). Both Mirenda et al. (2010) and Hooker et al. (2019) examined a one-, two-, three-, four-, five- and six-factor solution. All models were based either on conceptually existing models (e.g., Georgiades et al., 2010; Lam and Aman, 2007) or on present theories of RRBIs in the literature at the time. Results of Mirenda et al. (2010)'s CFA suggested that the five- and six-factor models yielded the best according to chi-square, RMSEA, AIC, and SRMR fit indices – as both models produced identical results. Hooker et al. (2019) yielded results that were mixed, though the five-factor model was ultimately favored as it produced minimally, slightly better-fit index estimates. For both Mirenda et al. (2010) and Hooker et al. (2019), the six-factor model was considered strongly, though the five-factor was chosen for the previously listed reasons, in addition to parsimony. Sturm et al. (2022)'s CFA included a one-through six-factor solution as well. Like the previous studies, the five- and six-factor models were suggested to be the best fit. However, the six-factor solution produced a slightly better fit. The five-factor solution was ultimately chosen based on interpretability and existing support from previous studies in the literature.

Conceivably, the most striking difference between the existing CFAs and the current study is the examination of multiple five-factor solutions. The Mirenda et al. (2010) and Hooker et al. (2019) studies listed their five-factor solutions as being similar to that of Lam and Aman (2007), though there were some differences in item inclusion and assignment. Sturm et al. (2022) conducted an EFA prior to their CFA, resulting in their own novel five-factor model. One strength was that none of the models listed had omitted items, allowing for direct comparison in this study. However, none of the existing CFAs that selected the five-factor model of the RBS-R examined the range of five-factor solutions that exist in the current literature. Therefore, study two allowed for a direct comparison of all existing five-factor models, including that of Mirenda et al. (2010) and Sturm et al. (2022).

To include a more thorough investigation of the frequently cited Lam and Aman (2007) five-factor model, the authors of the present study conducted an ancillary analysis involving the published factor model. By including a version of the Lam and Aman (2007) five-factor model, a theoretical comparison was allowed for and included for discussion. It should be noted that this reconstruction of the Lam and Aman (2007) model is not included in Chapter 4 of this study, as it was done purely for discussion purposes. The reconstruction involved an examination of the published pattern matrix available in Lam (2004)'s dissertation. All items that were originally assigned to factors in the Lam (2004) and later Lam and Aman (2007) study were kept on their factors. Omitted items from the model (e.g., items 21, 24, and 25) were then observed for their loadings across the five factors. In each case, each item was assigned to the factor that had the highest loading for that item. Though this is not an ideal representation of the Lam and Aman (2007) factor model, it provides an indication of the level of hypothetical fit for the highly regarded factor structure present in the literature. Appendix E includes a table of the original model (Table E1) and the reconstructed model (Table E2).

As discussed, AIC and BIC fit test results indicated that the study one EFA five-factor model was the best fitting model, compared to the other existing factor models from the current literature. With the introduction of the reconstructed Lam and Aman (2007) five-factor model, this was not entirely true. The AIC estimate for the reconstructed model was 20125.393 and BIC was 20163.933, boasting the lowest estimates for this fit measure across all factor models examined in the study. While this adds to the presence of support for a five-factor model overall, there remain questions about the item-level factor structure of the RBS-R, in this particular sample using special education staff raters. Though this was the first study to examine the instrument with this specific set of raters, the fact that Lam and Aman (2007)'s hypothetical model yielded a slightly better fit than the EFA-constructed model is noteworthy. While these results should not be considered fully, as they do not fully represent the published model, and are hypothetical, they do leave room for support of re-investigating the original Lam and Aman (2007) five-factor model for further examination. That is, the exclusion of previously omitted items may have a place in the use of the instrument when using teacher or special education staff raters over parent or caregiver raters. At best, the inclusion of these items allows for additional cross-model comparisons to be made in future studies, such as was accomplished in the present study.

Strengths

Research Contribution

The collective results of studies one and two delivered further insight into the existing clouded literature surrounding the RBS-R. Prior to the completion of this study, no researchers had examined the factor structure of the RBS-R with the use of teachers or special education staff as raters. The only prior study that conducted factor analyses of the RBS-R was conducted using a Spanish-translated version of the instrument (Martínez-González & Piqueras, 2017). Therefore, the study one EFA and the study two CFA were the first factor analytic studies of the RBS-R with non-caregiver raters using the original English version of the instrument.

The present studies' model yielded a novel factor structure that has not been observed in the literature. That is, one that includes a division of repetitive self-injurious behavior. The presence of the division of self-injury in both the five-factor and six-factor solutions generated by the study one EFA raises questions about teachers' and special education staff's perception of self-injury and body-focused repetitive behavior. There also remain questions about a potential model of repetitive self-injury that includes a higher-order, lower-order conceptualization. This exact conceptualization of self-injurious behavior is quite novel, though the presence of self-grooming behaviors themselves is not new. When considering RRBIs as a two-factor model (Georgiades et al., 2010; Cuccaro et al., 2003; Szatmari et al., 2006), self-injurious behavior itself typically falls into the broad lower-order category. Gold standard ASD diagnostic assessments such as the ADI-R (LeCouteur et al., 2013) place repetitive motor movements and self-injurious behaviors in the same grouping of lower-order categories as well. Studies that are more recent have found that body-focused repetitive behaviors in individuals with ASD may be due more to behavioral rigidity, anxiety, or cognitive inflexibility – all of which are considered "higher-order" processes. Posing this question about the conceptualization of repetitive behaviors, in general, is a strength in that it draws attention to the way the current literature considers repetitive behaviors – namely, self-injurious behaviors.

Sample Participants and Raters

One of the greatest strengths of studies one and two is that most information regarding participant characteristics (e.g., diagnosis, age, and cognitive ability) was readily available for inclusion in the studies. Past studies (Bishop et al., 2013; Lam & Aman, 2007; Russell et al., 2019) utilized wide-scale research agencies or existing databases to recruit participants for their studies. Therefore, participant demographics were at times limited, regarding specific, but important characteristics.

It is important to highlight not only the uniqueness of the sample raters used in this study but also the demographics of participants included in studies one and two. Various existing studies that used caregiver ratings have received censure for their use of a limited age range. Mirenda et al. (2010) underscored Lam and Aman (2007)'s use of a sample with primarily older individuals – including a sample ranging in age from 16 to 56 years old. Interestingly, Mirenda et al. (2010) and Hooker et al. (2019) utilized samples with limited age ranges, with individuals ranging in age from four to nine years old, and two to five years old, respectively. Other studies (Bishop et al., 2013; Russell et al., 2019; Sturm et al., 2022) using caregiver ratings included a broader age range – four to 18 years old. It should be noted that Martínez-González and Piqueras (2017) included teacher ratings with a similar sample size to that used in studies one and two – including participants ranging in age from three to 63 years old. However, this study was conducted using the Spanish-translated version of the RBS-R, unlike the current study, which used the original RBS-R in English. The current study utilized a large age range – ranging from three to 21 years old, while also including raters from special education staff, rather than caregivers. This provided a novel insight into the validity of the RBS-R using a unique subset of sample demographics and raters.

This study has the most gender-diverse sample, compared to other EFA and CFA studies of the RBS-R. Study one had a sample with 78.2% male participants and study two had a sample with 77.25% male. Existing studies included a range of 80.9% (Sturm et al., 2022) to 86.6%

(Russell et al., 2019) males. This information is helpful, as the gender-focused ratios in studies one and two for participants are closest to the most recent male-to-female prevalence rate for ASD, which is 3:38 to 1 (Maenner et al., 2023). Therefore, the results of this dissertation may provide accurate information about the factor structure of the RBS-R in a sample of children with ASD, considering gender differences. It should be noted that this is not true for the samples' representations of racial and ethnic demographics, which will be discussed in the limitations of this study.

All participants were diagnosed by licensed clinicians using either the DSM-IV or DSM-5 criteria or by special education eligibility criteria. This clarification of diagnostic confirmation proved a strength, as several of the existing studies in the literature relied on caregiver reports (e.g., Bodfish & Lewis, 2002; Lam & Aman, 2007) to confirm diagnoses. Other studies obtained their diagnostic information from existing wide-scale research databases (Bishop et al., 2013; Russell et al., 2019; Sturm et al., 2022), only listing that participants had an existing diagnosis of ASD. In contrast, all participants included in studies one and two had a confirmed diagnosis by clinical staff who had strong experience in the field of ASD and developmental disabilities. Further, all ratings of the RBS-R were completed by raters who were familiar with the participants. The special education agency from which the data was obtained allowed for a near 1:1 staff rater-to-student ratio, meaning that each staff member likely only rated one participant each. Allowing for such a strong level of rater independence was assumed to hinder subgroups of students from becoming grouped into nests from ratings made by one staff member. This differs from the one study (Martínez-González & Piqueras, 2017) that utilized teacher ratings, as all ratings were listed as being completed with raters who were familiar with ASD itself, rather than the participants who were being rated. The specific knowledge of diagnoses for each participant, as well as the unique rater-specific knowledge of each participant, was a strength, as no other EFA or CFA of the RBS-R has accounted for such issues in their samples.

Methodology

Across both studies one and two, use of the most updated best practices for EFA and CFA methodology was used. While several of the existing studies utilized acceptable methods, especially for more modern CFA studies, the present research built upon existing studies as well as standards set for best practices (e.g., Basto & Pereira, 2012; Byrne, 2012; Osborne & Banjanovic, 2016) to provide the most statistically sound results. Study one boasted a strength in the level of detail provided throughout each step of the EFA process overall. The most level of detail regarding methodology among the RBS-R EFA literature was expected to stem from Lam (2004), as this was the dissertation that came before the later published Lam and Aman (2007) article. However, research from Lam (2004) and later Lam and Aman (2007) were severely lacking in important information. That is, study one provided much more specific information regarding item-level descriptive statistics – as Lam (2004) or any other existing EFA studies (e.g., Bishop et al., 2013; Russell et al., 2019) did not include such information.

Regarding the EFA criterion used for extraction, study one used a rigorous combination of methods that have not been observed across the current literature. The use of the inter-item polychoric correlation matrix for this study was a strength, as the use of Pearson correlation matrices often underestimates correlations when used with data from ordinal scales such as the RBS-R. Study one also included the use of principal axis factoring (PAF), which is highly preferred for the sample included in the present study. Study one also used an oblique rotation for extraction. However, the existing studies utilized methods such as Principal Components Analysis (PCA; Lewis & Bodfish, 2002), Ordinary Least Squares with oblique quartimax rotation (Lam, 2004; Lam & Aman, 2007), and EFA with promax rotation (Bishop et al., 2013). The closest use of EFA methodology to the present study was observed by Russell et al. (2019), which included an EFA using PAF and a promax rotation. The use of the PAF in both study one and in Russell et al. (2019) is highlighted, as this is a significant consideration for the non-normal data that is often observed from ASD samples. This procedure is also often preferred for such samples as it is more robust for samples with expected deviations from data normality.

Study one retention criteria was also a strength, as the application of several strategies occurred across a range of indices. The use of indices such as the Guttman-Kaiser Criterion, parallel analysis, and the MAP test were included in the study. Further, factor solutions were examined for interpretability by a team of independent ASD-focused researchers who named the potential factors according to the strongest constructs present across each factor and solution. Following independent considerations, the team convened and determined the most interpretable factor and settled on factor names. To the author's knowledge, no existing studies in the RBS-R EFA literature used such rigorous methods for factor consideration, interpretability, and decision-making. As such, study one utilized stronger, more accurate methods to guide factor retention decision-making, when compared to the existing studies in the present RBS-R EFA literature with ASD samples.

Like study one, the methodology used in study two's CFA followed guidelines for best practice and included detailed information at each step. The use of the Mplus Weighted Least Squares Mean-Variance (WLSMV) estimator was a strength, as it addressed the ordinal RBS-R data, in addition to the non-normal data existing in the ASD sample. The WLMSV estimator is highly regarded for ordinal data and is robust when considering normality deviations (DiStefano & Morgan, 2014). Only one of the three published RBS-R CFA studies (Hooker et al., 2019) utilized a WLSMV estimator, leaving vulnerabilities regarding the use of the methodology for the data used in other studies.

Again, like study one, study two also included a wide range of fit indices as part of the CFA methodology. Indices such as χ^2, SRMR, RMSEA, CFI, and TLI were used to address model fit within the WLSMV Mplus estimator. All the listed model fit indices were included in some way across all existing RBS-R CFA studies, though no study itself has included them all at once. That is, Mirenda et al. (2010) only utilized χ^2, RMSEA, CFI, and TLI; Hooker et al. (2019) only utilized RMSEA, CFI, and TLI. Sturm et al. (2022) utilized RMSEA, SRMR, CFI, and TLI. However, it should be noticed that Sturm et al. (2022) utilized a CFA using a MLE estimator, instead of a WLSMV estimator. To add to its strengths, study two also used the AIC and BIC to compare non-nested models using the Robust Maximum Likelihood (MLR) estimator in Mplus, comparing the fit of all existing models within the validation sample that utilized the same inter-item covariance matrix.

Model Testing

One of the greatest strengths of the present research is the inclusion of all existing RBS-R factor models to date. In the present study, six different factor models were available for testing in the CFA: (a) the published six-factor model from the RBS-R (Bodfish et al., 2000), (b) the five-factor model from Lam (2004) and Lam and Aman (2007), (c) the five-factor model from Bishop et al. (2013), (d) the four-factor model from Russell et al. (2019), (e) the Mirenda et al. (2010) five-factor model, (f) the six-factor model from Martínez-González and Piqueras (2017), and (g) the five-factor model from Sturm et al. (2022). At the time of these studies, no published CFA studies have examined all the existing factor solutions previously listed within their study samples. All three published CFA studies (Hooker et al., 2019; Mirenda et al., 2010; Sturm et al.,

2022) examined the model fit for one- through six-factor models – all of which were derived from their own EFA or based on theories of RRBIs in the literature. This study provided a comprehensive examination and comparison of all RBS-R factor models to date, using the CFA validation sample. Of note, these models were tested using ratings made by special education staff, while the existing studies were conducted using caregiver raters.

With this, results from the CFA in study two provide novel information for discussing the differentiation between the use of teacher/special education staff and caregiver raters for the RBS-R. The results from study two provided evidence that a five-factor solution may be an appropriate fit for special education staff raters in an ASD population, though future research targeting replication of results is still needed. When looking specifically at the AIC and BIC data from various models, the models utilizing a five-factor model stood out as the best fit overall in this validation sample. While the study one EFA five-factor model stood out as the best fitting model compared to all other models included for comparison, Sturm et al. (2022)'s model was also highlighted as a solid model. The reconstruction of the Lam and Aman (2007) five-factor model for ancillary purposes also posed some questions about which five-factor model may be the best fit for teacher and special education staff raters in this sample. Even further though, the inclusion of this reconstructed model is a strength, as it allowed for hypothetical consideration of this model for discussion purposes – although it should be noted that this factor model cannot be included in any meaningful discussion of the factor models, as it is only a hypothetical version of the published model. However, the consistency of results that supported the study one EFA five-factor model provides evidence of the strength of the model for this sample of special education staff raters overall. Regardless of the Lam and Aman (2007) five-factor model, the results of the study two CFA leave room for future research to be completed in this domain, but the contribution of this research to the existing literature is a noteworthy strength.

Limitations

Sample Demographics and Methods

Despite the several strengths presented regarding sample demographics, there were also some limitations to be noted. Although the individuals in the sample all received diagnoses by licensed psychologists under either the DSM-IV, DSM-5 or the special education criteria, there were no requirements for the use of gold-standard instruments such as the ADI-R or ADOS-2 for diagnostic purposes. It is likely that most participants did receive a diagnosis using these tools; however, it was not necessarily a requirement from the special education agency. This goes against best practice standards set for ASD diagnoses, which is complex and calls for the inclusion of a wide range of assessment tools (e.g., structured observations, rating scales), and the use of the ADI-R and ADOS-2 (Aiello et al., 2017; Esler & Ruble, 2015; Martin et al., 2018). Although diagnoses and special education eligibility were required to be made by licensed clinicians, there was no set assessment battery to be used for each participant in either sample for studies one or two. This is another limitation of the study, as a large variety of cognitive assessments to derive the intelligence quotients for each participant. The most updated cognitive assessment scores were included in this study for each sample case, but there was no set battery or assessment used for all participants. That is, over 20 different assessments were used to determine participants' cognitive functioning across the two studies. A large number of assessments were conducted to provide participants with the most accurate assessment of their abilities, considering factors such as age and verbal language abilities.

While the samples in studies one and two appeared to reflect accurately the most current demographics and prevalence rates for ASD regarding gender representation, this was not true for racial and ethnic representation. In both samples, White participants were overrepresented in the samples for studies one and two, as they made up 73.50% (study one) to 78.11% (study two) of participants. The most current prevalence data suggest that ASD diagnoses for White individuals are lower when compared to the collective whole of non-White (e.g., Black, Hispanic, Asian, Native American, and Pacific Islander) individuals (Maenner et al., 2023). There is some research to suggest that children of diverse racial and ethnic backgrounds (e.g., Black, Filipino, Vietnamese) are at a higher risk for receiving an ASD diagnosis when compared to White children (Becerra et al., 2014). Although, it must be stated that minority status does not predict an ASD diagnosis over a non-ASD diagnosis (Cuccaro et al. 2007; Herlihy et al., 2014). This information might suggest issues with the most current depiction of racial and ethnic differences in the presence of ASD prevalence rates. However, it should not be understated that White participants were vastly overrepresented in studies one and two, nonetheless.

Extraction Criteria

Though it is beyond the scope of the present studies to outline and describe the complexities of the various types of factor rotations used in the EFA literature, it is important to note that this study utilized a different rotation method than all others in the literature. Study one included the use of a direct oblimin rotation, while all other studies in the current literature used promax and varimax rotations. The differences in methods used in the EFA pose challenges for the comparison of solutions and methods across the RBS-R EFA studies. While was is not feasible nor recommended to run every possible factor rotation method on the sample data for this study, the inclusion of similar methods may have allowed for more likenesses between existing studies and the present research. Future studies may look to include a broader exploration of appropriate factor rotation methods for such comparisons.

While various factor extraction methods were used in study one, the inclusion of such extraction criteria may also present a limitation. While study one included four different extraction methods (e.g., scree test, Guttman-Kaiser criterion, parallel analysis, and MAP test), no other published RBS-R EFA studies included the MAP test as a method. While this was listed as a strength previously, including this extraction method posed challenges in determining the strength of the chosen factor structure in this study. It should not be taken that use of the MAP test itself is a limitation. Generally, the MAP test is the most robust extraction technique in the current EFA literature (Osborne & Banjanovic, 2016). Whereas other extraction methods (e.g., scree test, Guttman-Kaiser criterion) are more subjective in nature. However, because other studies did not use the MAP test, nor parallel analysis, in their analyses to guide factor retention, it is difficult to directly compare the outcomes of the EFA in study one to existing studies. This opens the discussion for hypothesizing potentially different factor solutions in existing studies, had authors used a more robust approach such as the MAP test.

Sample Size

Even though the sample sizes for both studies one and two were adequate for the EFA and CFA procedures, larger sample sizes for each study would have been ideal. Study one included 234 participants and study two included 233 participants – both meeting the MacCallum et al. (1999) ratio of variables to factors for the RBS-R, calling for a sample size between 100 to 200 participants. A strict sample size requirement is not present in the EFA/CFA literature. However, there is some agreement that at least 200 participants are ideal (Floyd & Widamin, 1995) to be theoretically sufficient in supporting an EFA factor solution. For CFA, there are even fewer guidelines regarding sample size. There is consensus that larger sample sizes tend to support the conduction of a CFA and yield stable results (Harrington, 2009; MacCallum et al., 1999; Marsh et al., 2009; Schmitt, 2011) and MacCallum calls for the same sample size conceptualization for CFA and is present for an EFA. However, there is an assumption that non-normally distributed data calls for increased sample size – changing from 150 in a normal distribution to 265 for a non-normal distribution of data (Muthén & Muthén, 2002). There were no concerns about the adequacy of the sample size for either study. Despite this fact, a larger sample size for both studies one and two may have provided more stable results. That is, results may have been less susceptible to estimates and standard errors.

It is also important to point out that when compared to the current literature, studies one and two present the lowest sample size across all RBS-R EFA and CFA studies. In the EFA literature, sample sizes range from 350 participant cases (Lam & Aman, 2007) to 2,093 participant cases (Russell et al., 2019). In the CFA literature, sample sizes range from 287 participant cases (Mirenda et al., 2010) to 15,318 participant cases (Sturm et al., 2022). Again, the sample sizes for the present studies do not take away from the strength of the results by any means. On the other hand, including more participants would likely make a stronger case for the factor model produced in the EFA, as well as the fit indices reported in the CFA.

Generalizability

Although the present study provides novel information regarding the validity of the RBS-R by use of special education staff raters, there exists a limitation in the generalizability of the outcomes. With this being the first study to examine the published RBS-R through the lens of special education staff raters, it is not clear if any differences in the results of the current studies are generalizable to other types of raters. While this study found a five-factor structure, there were differences in the number of items included in the model, as well as in the names of some factors, compared to existing five-factor models from caregiver raters. It should be noted that the only study that used teacher raters to examine the validity of the Spanish version of the RBS-R found a six-factor structure. Differences in structure for that particular study could potentially be due to language or cultural differences for the Spanish-speaking population, though it is beyond the scope of this study to explore such a hypothesis.

One limitation of the present study was that all items were included in the final factor model. In other studies, several items were omitted, either due to no items loading uniquely on a factor, low interpretability, or due to no significant item-to-factor loadings. The study one EFA suggested minimal substantive cross-loadings (four items – items 24, 29, 34, and 41) and all items appeared to clearly fit the factors that they were assigned to. While this is a strong characteristic of the model, it is a limitation in that this led to fewer comparisons being made in the CFA. That is, studies such as Bishop et al. (2013), Lam and Aman (2007), and Russell et al. (2019) all removed at least three items, if not more, from their models. While some ancillary analyses of the reconstructed Lam and Aman (2007) five-factor model to hypothetically structure factors to include omitted items by placing the highest-loading items onto their respective factors took place, there are serious compromises to such factor interpretations. However, this leaves questions about the truest, "best fit" model for special education staff raters in a sample of children with ASD, as the reconstructed model posed even better fitting results than the study one EFA. However, with fewer comparisons to be examined, it is difficult to say that the model generated in the EFA is a better fit than any other existing models that were left out of the AIC and BIC direct model comparisons.

Implications

With the use of special education staff as raters on the RBS-R, there lie several implications of the present study across theoretical, practical, and research considerations. Beyond this unique strength, results yielded from studies one and two contribute important information and suggestions for future use and examination. As this dissertation is the first study, to the author's knowledge, to include special education staff in consideration of the factor structure of the RBS-R, it is important to note that future studies should continue to follow suit to provide stronger evidence of the generalizability of results.

Theoretical

Broadly, the results of the EFA in study one provided strong evidence for a five-factor solution, consistent with previous studies examining the RBS-R using caregiver ratings of individuals with ASD. When considering the factor structure of the instrument overall, there is overwhelming support across existing studies (e.g., Bishop et al., 2013; Hooker et al., 2019; Lam & Aman, 2007; Mirenda et al., 2010; Sturm et al., 2022) to support a five-factor conceptualization of the RBS-R – regardless of the raters used. The CFA in study two provided a confirmation of the fit of the five-factor model, as only the five-factor models (e.g., Mirenda et al., 2010; Sturm et al., 2022; study-one EFA model) yielded results indicating the best fit according to AIC and BIC values. It is remarkable that the best-fitting five-factor models all yielded slightly different models – namely, the study one EFA with a unique Body-Focused Repetitive Behaviors factor. This perhaps provides additional support for the five-factor conceptualization of RRBs, at least on the RBS-R that appeared to exist in the literature amongst caregiver raters. Study one and two results differed from those yielded by Martínez-González and Piqueras (2017), who used teacher ratings as well. However, it is hard to strongly argue measures of comparison between the two studies, as Martínez-González and Piqueras (2017) used a Spanish-translated version of the instrument in a Spanish-speaking population. Aside from the potential language translation issues, there also remain questions about cultural differences in reporting and understanding of RRBIs overall. In any case, the present study raises key theoretical implications in the conceptualization of repetitive behaviors from a five-factor orientation.

Another important theoretical implication that stems from the results of studies one and two comes from the analysis of the inter-factor correlations for the present study's five-factor model. The study one inter-factor correlation matrix yielded some strong correlations between factors – mainly between the Insistence on Sameness factor and Stereotyped Behaviors factor. The only factor that did not include any strong correlations was the Body-Focused Repetitive Behaviors factor (correlations with other factors ranging from .10 to .22). This is important to consider because this factor is a division of the original Self-Injurious Behavior factor, which it only yielded a correlation of .22. The study two CFA inter-factor correlation matrix for the five-factor solution with the validation sample yielded somewhat better results – with ranges of correlations ranging from .29 to .55 – the highest being with Self-Injurious Behaviors. The range of inter-factor correlations for this factor poses theoretical questions about its presence on the instrument overall. Even further, the existence of this factor itself was unanticipated, as no other studies have found evidence for such a factor – leaving no clear path for comparison in existing studies. This factor should be explored in future research, as it presents a distinctive conceptualization of self-injurious behavior and repetitive behaviors.

The dissolution of the Restricted Interests factor for this study is another outstanding theoretical implication. Broadly, the construct measured by the RBS-R is repetitive behaviors and was developed (and has been used for) for use with individuals with ASD. As such, the inclusion of restricted interest-related behaviors on the original RBS-R is fitting regarding face validity. The DSM-IV-TR (APA, 2000) and DSM-5 (APA, 2013) both clearly include restricted behaviors as a core feature of ASD. With this information, there is no question about the inclusion of items relating to restricted interests and behaviors on the RBS-R, as they hold a very meaningful place as per diagnostic criteria. However, there is fairly strong evidence to suggest that this construct is not well-represented on the instrument overall. In all previous RBS-R factor analytic studies (using caregiver raters), all factors representing restricted behavior have been proven to be the weakest factor – with most studies presenting this factor with four items or fewer. It is important to consider the exemplification of this construct at the item level on the RBS-R overall. With this, revisions of the items included on the instrument may be helpful to better capture a core diagnostic feature of ASD.

Practical

Though the RBS-R has consistently been validated for use with caregiver raters, there are several opportunities for practitioners to use the instrument in treatment and school settings. An overwhelming number of studies have used the RBS-R as an outcome measure for treatments and interventions for research purposes, as well as an assessment measure in various settings (Rojahn et al., 2013; Schertz et al., 2016; Wolff et al., 2016). With the contribution of the results from studies one and two, there may be some evidence to support the RBS-R to special education staff as a measure of an outcome as well. In addition, staff who work closely with individuals receiving treatment via Applied Behavior Analysis (ABA) may provide helpful assessment information regarding the clients they work with using the RBS-R – though it should be noted that the RBS-R cannot be used with standardized norms or used solely for assessment purposes.

It should also be noted that the scope of application for the RBS-R is limited in that the tool is often used as an outcome measure for research studies. While some practitioners have used the tool for progress monitoring, it has not been established as a core assessment tool for individuals with ASD. The results of studies one and two contribute unique information regarding the overall factor structure of the RBS-R with school staff raters, it does not necessarily inform the use of the tool in a practical setting. The RBS-R was developed as a research instrument, though there is an abundance of citations reporting the need for a tool of its kind in the field of ASD intervention and assessment. However, the lack of standardized scores and norms for comparison against others of the same age poses problems in this realm of practice. Perhaps a rather important recommendation for future research is to establish norms to allow for the expansion of the RBS-R. Until this is completed, the RBS-R must be used with caution and not in isolation to algin with best practices of using multisource and methods for ASD-related assessments. With the present EFA and CFA findings, there is evidence to support the internal structure of the RBS-R for use with special education staff raters for children with ASD, although results are truly limited to application amongst the studies' samples. While there is still room for future studies to continue this research regarding generalization, these initial findings contribute to the support of RBS-R use in practical settings with education staff raters.

One important consideration that arises from the results of these studies is observed in the strength of factors. In all RBS-R factor analytic studies, including the present studies, the Stereotyped Behaviors factor has repeatedly emerged as the strongest, clearest factor on the RBS-R. This was true for caregiver raters, as well as for special education staff raters as well. Of note for practitioners, stereotyped behaviors have been demonstrated to interfere with observational learning (Varni et al., 1979) and play skills instruction (Koegel et al., 1974). Such behaviors have also been suggested to negatively affect the reception of auditory stimuli (Lovaas et al., 1977; Kanakri et al., 2017) and the completion of discrimination tasks (Koegel & Covert, 1972). This information is particularly useful for special education staff, as these behaviors appear to cause significant intrusion upon functioning crucial for both educational instruction and skills-based learning. Given the interference that these behaviors may inflict on academic performance, the use of the RBS-R, especially the Stereotyped Behaviors factor, may be helpful in better understanding specific student needs regarding accommodations.

An additional consideration for practitioners to reflect at the item level lies within a few items that fall under the present study's Compulsive Behaviors factor. Items 23, 24, and 25 all inquire about the individual's routines surrounding eating/mealtime, sleeping/bedtime, and self-care/bathroom and dressing behaviors. In the present study using special education staff raters, these three items provided the highest number of missing ratings from raters (though this was not significant enough to warrant removal from the overall model as only three cases from study one and two from study two had left these items incomplete). Anecdotally, these raters sometimes wrote in a question mark by these items, making note of their uncertainty about student behavior at home. Depending on the school setting, practitioners likely have access to information regarding their student's behaviors during these routines. For the present study, this was not a significant issue in rater response, as the special education agency where the data was collected provides a range of services – some of which include opportunities for mealtime and self-care or bathroom and dressing routines to occur. However, it is recommended that practitioners who use the RBS-R consider their ability to rate students on these items, especially if students do not engage in such routines in the rater's presence on a regular basis. From another perspective, retaining these items for raters who *do* observe such behaviors is undoubtedly helpful for measures of treatment outcomes or suggestions of potential student accommodations.

Research and Research Recommendations

With the emergence of the data yielded from studies one and two, results leave several opportunities for future research of the RBS-R with teacher or special education staff raters of children with ASD. Future studies may build on the limitations set out across both studies one and two, as well as contribute to this novel research area of examining RBS-R raters who are not caregivers. Given that this was the first study to examine the use of school-based raters using the original English version of the RBS-R, there were limited existing psychometric research studies for comparison of validity. As a broad area of research for future studies, it is recommended strongly that future studies continue to examine the validity and factor structure of this instrument with teacher or school staff raters using the originally published RBS-R.

Given the differences present in the present studies' EFA and CFA results compared to existing studies, performing additional analyses on the RBS-R with teacher or special education staff raters would certainly contribute to the current literature. The study one EFA yielded no minimal cross-loadings from items to factors on the chosen five-factor solution. However, several studies (e.g., Bishop et al., 2013; Lam and Aman, 2007; Mirenda et al., 2010; Russell et al., 2019) using caregiver raters reported cross-loadings on their items – some so high that items were omitted from the final factor model reported in studies. Some studies also removed items due to a lack of any unique substantive item loadings on factors. Studies that removed items generally shared the same items. However, the present study did not flag these items as problematic when using special education staff raters. Such items should be examined for cross-loadings (or no substantive loadings) using education staff raters to further contribute to understanding the instrument when used with such raters. Further, the presence of the new factor, Body-Focused Repetitive Behaviors should be examined further in future studies as well. The presence of this factor suggests a type of higher-order, lower-order conceptualization of self-injurious behavior that has emerged in any studies to date. This may be a result of differences in raters used, due to participant demographic characteristics (e.g., language or cognitive abilities), or another potential moderating factor. Regardless, examination in future studies should consider these speculations and observe any differences in item placement or item loading that effects interpretation.

Results yielded from the study two CFA provide another avenue for future research as well. While the study one EFA five-factor model was distinguished as the best-fitting model to the CFA validation sample, there were other models that also yielded noteworthy results as well. Namely, the Sturm et al. (2022) five-factor model, Mirenda et al. (2010) five-factor model, and the reconstructed, hypothetical Lam and Aman (2007) five-factor model also presented as potential fits for the sample as rated by special education staff. Because this was the first study to compare comprehensively all RBS-R factor models present in the literature, there are no other data to compare study two results against for evaluation. Future RBS-R CFA studies should seek to include such comparisons -- either using caregivers or school staff as raters, to contribute more data for comparison of such results across a myriad of participants and rater types.

One aspect of the study that was not considered in the analysis was the potential moderator of cognitive ability and language functioning on the factor structure of the RBS-R. While each participant had a report of their cognitive and/or language abilities, consideration of their effects on the way repetitive behavior manifests in individuals may provide helpful information. Future research should look further into the application of the RBS-R factor structure in samples of individuals with ASD in specific subsets of language and cognitive abilities. The present study included a broad range of these demographic characteristics but did not include a specific examination of cognitive and language abilities as moderators on the RBS-R factor structure. Future studies would benefit by considering this limitation to better understand how repetitive behaviors may present differently for individuals who fall in different ranges of the autism spectrum. In a similar vein, future studies may also consider the DSM-5 severity levels for repetitive behaviors associated with an ASD diagnosis. That is, there may be helpful information regarding the factor structure of the RBS-R and how it conceptualizes repetitive behaviors when considering the varying severity levels of ASD (e.g., Level 1 – Requiring Support, Level 2 – Requiring Substantial Support, Level 3 – Requiring Very Substantial Support). Such information may provide a more comprehensive understanding of potential moderating variables.

REFERENCES

Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: implications of cross-informant correlations for situational specificity. *Psychological Bulletin, 101*(2), 213-232. https://doi.org/10.1037/0033-2909.101.2.213

Achenbach, T. M., & Rescorla, L. A. (2000). *Manual for the ASEBA preschool forms and profiles* (Vol. 30). University of Vermont, Research Center for Children, Youth, & Families.

Aiello, R., Ruble, L., & Esler, A. (2017). National study of school psychologists' use of evidence-based assessment in autism spectrum disorder. *Journal of Applied School Psychology, 33*(1), 67-88. https://doi.org/10.1080/15377903.2016.1236307

Akaike, H. (1987). Factor analysis and AIC. *Psychometrika*, *52*(3), 317-332. https://doi.org/10.1007/bf02294359

Alavi, M., Visentin, D. C., Thapa, D. K., Hunt, G. E., Watson, R., & Cleary, M. (2020). Chi-square for model fit in confirmatory factor analysis. *Journal of Advanced Nursing*, 76(9), 2209– 2211. https://doi.org/10.1111/jan.14399

Allison, P. D. (2002). *Quantitative applications in the social sciences: Missing data.* Sage Publications Ltd. https://doi.org/10.4135/9781412985079

Aman, M. G., Singh, N. N., Stewart, A. W., & Field, C. J. (1985). The Aberrant Behavior Checklist: A behavior rating scale for the assessment of treatment effects. *American Journal of Mental Deficiency*, 89(5), 485-491.

American Psychiatric Association. (2000). Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR). https://doi.org/10.1176/appi.books.9780890423349

American Psychiatric Association. (2013). Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5). https://doi.org/10.1176/appi.books.978089042559

Arnott, B., McConachie, H., Meins, E., Fernyhough, C., Le Couteur, A., Turner, M., Parkinson, K., Lucia, V., Vittorini, L., & Leekam, S. (2010). The Frequency of Restricted and Repetitive Behaviors in a Community Sample of 15-Month-Old Infants. *Journal of Developmental & Behavioral Pediatrics, 31*(3), 223-229. https://doi.org/10.1097/dbp.0b013e3181d5a2ad

Asperger, H. (1944). Die "Autistischen Psychopathen" im Kindesalter. *Archiv für Psychiatrie und Nervenkrankheiten*, *117*(1), 76-136. https://doi.org/10.1007/bf01837709

Autism and Developmental Disabilities Monitoring Network Surveillance Year 2002 Principal Investigators: CDC. (2007). Prevalence of Autism Spectrum Disorders --- Autism and

Developmental Disabilities Monitoring Network, 14 Sites, United States, 2002. *Morbidity and Mortality Weekly Report: Surveillance Summaries, 56* (No. SS-1).

Azad, G., Sridhar, A., Taormina, I., & Roter, D. L. (2023). Parent–teacher communication for children on the autism spectrum: An examination of communication skill use and problematic communication. *Psychology in the Schools*, *60*(3), 691-706. https://doi.org/10.1002/pits.22787

Baglin, J. (2014). Improving your exploratory factor analysis for ordinal data: A demonstration using FACTOR. *Practical Assessment, Research, and Evaluation*, *19*(1), 5. https://doi.org/10.7275/dsep-4220

Baio, J., Wiggins, L., Christensen, D. L., Maenner, M. J., Daniels, J., Warren, Z., Kurzius-Spencer, M., Zahorodny, W., Rosenberg, C. R., White, T., Durkin, M. S., Imm, P., Nikolaou, L., Yeargin-Allsopp, M., Lee, L., Harrington, R., Lopez, M., Fitzgerald, R. T., Hewitt, A., … Dowling, N.F. (2018). Prevalence of autism spectrum disorder among children aged 8 years - autism and developmental disabilities monitoring network. *Morbidity and Mortality Weekly Report (MMWR) Surveillance Summaries*, *67*(6), 1–23. https://doi.org/10.15585/mmwr.ss6706a1.

Barber, A. B., Wetherby, A. M., & Chambers, N. W. (2012). Brief report: Repetitive behaviors in young children with autism spectrum disorder and developmentally similar peers: A follow up to Watt et al. (2008). *Journal of Autism and Developmental Disorders*, *42*(9), 2006-2012. https://doi.org/10.1007/s10803-011-1434-3

Bartlett, M. S. (1950). Tests of significance in factor analysis. *British Journal of Statistical Psychology*, *3*(2), 77-85. https://doi.org/10.1111/j.2044-8317.1950.tb00285.x

Basto, M., & Pereira, J. M. (2012). An SPSS R-menu for ordinal factor analysis. *Journal of Statistical Software*, *46*, 1-29. https://doi.org/10.18637/jss.v046.i04

Batista-Foguet, J. M., Saris, W., Boyatzis, R. E., Serlavós, R., & Velasco Moreno, F. (2019). Multisource Assessment for Development Purposes: Revisiting the Methodology of Data Analysis. *Frontiers in psychology*, *9*, 2646. https://doi.org/10.3389/fpsyg.2018.02646

Beavers, A. S., Lounsbury, J. W., Richards, J. K., Huck, S. W., Skolits, G. J., & Esquivel, S. L. (2013). Practical considerations for using exploratory factor analysis in educational research. Practical Assessment, Research & Evaluation, 18(6). https://doi.org/10.7275/qv2q-rk76

Becerra, T. A., von Ehrenstein, O. S., Heck, J. E., Olsen, J., Arah, O. A., Jeste, S. S., Shafali, S. J., Rodriguez, M., & Ritz, B. (2014). Autism spectrum disorders and race, ethnicity, and nativity: a population-based study. *Pediatrics*, *134*(1), e63-e71. https://doi.org/10.1542/peds.2013-3928

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*(2), 238. https://doi.org/10.1037/0033-2909.107.2.238

Birnbaum, R. (2020). *Exploratory and Confirmatory Factor Analysis of the Aberrant Behavior Checklist-Community in an Autism Spectrum Disorder Sample with Ratings Completed by Special Education Staff*. Michigan State University.

Bishop, S. L., Hus, V., Duncan, A., Huerta, M., Gotham, K., Pickles, A., Kreijer, A., Buja, A., Lund, S., & Lord, C. (2013). Subcategories of restricted and repetitive behaviors in children with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 43*(6), 1287-1297. https://doi.org/10.1007/s10803-012-1671-0

Bishop, S. L., Reichler, J., & Lord, C. (2006). Association between restricted and repetitive behaviors and nonverbal IQ in children with autism spectrum disorders. *Child Neuropsychology, 12*, 247–267. https://doi.org/10.1080/09297040600630288

Bitsika, V., Sharpley, C. F., Andronicos, N. M., & Agnew, L. L. (2016). Prevalence, structure and correlates of anxiety-depression in boys with an autism spectrum disorder. *Research in Developmental Disabilities*, *49-50*, 302-311. https://doi.org/10.1016/j.ridd.2015.11.011

Bodfish, J. W., Crawford, T. W., Powell, S. B., & Parker, D. E. (1995). Compulsions in adults with mental retardation: Prevalence, phenomenology, and comorbidity with stereotypy and self-injury. *American Journal on Mental Retardation*.

Bodfish, J. W., Symons, F. W., & Lewis, M. H. (1999). The Repetitive Behavior Scale. Western Carolina Center Research Reports.

Bodfish, J. W., Symons, F. J., Parker, D. E., & Lewis, M. H. (2000). Varieties of repetitive behavior in autism: Comparisons to mental retardation. *Journal of Autism and Developmental Disorders*, *30*(3), 237-243. https://doi.org/10.1023/A:1005596502855

Bolte, E. E., & Diehl, J. J. (2013). Measurement tools and target symptoms/skills used to assess treatment response for individuals with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 43*(11), 2491-2501. https://doi.org/10.1007/s10803-013-1798-7

Bölte, S., Duketis, E., Poustka, F., & Holtmann, M. (2011). Sex differences in cognitive domains and their clinical correlates in higher-functioning autism spectrum disorders. *Autism*, *15*(4), 497-511. https://doi.org/10.1177/1362361310391116

Boyd, B. A., McDonough, S. G., & Bodfish, J. W. (2012). Evidence-based behavioral interventions for repetitive behaviors in autism. *Journal of Autism and Developmental Disorders*, *42*(6), 1236-1248. https://doi.org/10.1007/s10803-011-1284-z

Boyd, B. A., Woodard, C. R., & Bodfish, J. W. (2011). Modified exposure and response prevention to treat the repetitive behaviors of a child with autism: A case report. *Case Reports in Psychiatry*, *2011*. https://doi.org/10.1155/2011/241095

Brinkley, J., Nations, L., Abramson, R. K., Hall, A., Wright, H. H., Gabriels, R., Gilbert, J. R., Pericak-Vance, M. A. O.., & Cuccaro, M. L. (2007). Factor analysis of the Aberrant Behavior Checklist in individuals with autism spectrum disorders. *Journal of Autism and*

Developmental Disorders, 37(10), 1949-1959. https://doi.org/10.1007/s10803-006-0327-3

Bristol, M. M., Cohen, D. J., Costello, E. J., Denckla, M., Eckberg, T. J., Kallen, R., Kraemer, H. C., Lord, C., Maurer, R., McIlvane ,W. J., Minshew, N., Sigman, M., Spence, M. A. (1996). State of the science in autism: Report to the national institutes of health. *Journal of Autism and Developmental Disorders*, *26*(2), 121-154. https://doi.org/10.1007/bf02172002

Brown, T. A. (2006). Confirmatory factor analysis for applied research. Guilford Press Publications.

Brown, J. D., Wissow, L. S., Gadomski, A., Zachary, C., Bartlett, E., & Horn, I. (2006). Parent and teacher mental health ratings of children using primary-care services: interrater agreement and implications for mental health screening. *Ambulatory Pediatrics: The Official Journal of the Ambulatory Pediatric Association*, *6*(6), 347–351. https://doi.org/10.1016/j.ambp.2006.09.004

Brown, T. A. (2015). *Confirmatory factor analysis for applied research*. Guilford publications.

Brown, T. A., & Moore, M. T. (2012). Confirmatory factor analysis. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 361-379). The Guilford Press.

Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods and Research*, *21*(2), 230-258. https://doi.org/10.1177/0049124192021002005

Bull, C. R., Bull, R. M., & Rastin, B. C. (1992). On the sensitivity of the chi-square test and its consequences. *Measurement Science and Technology*, *3*(9), 789-795. https://doi.org/10.1088/0957-0233/3/9/001

Byrne, B. M. (2012). *A primer of LISREL: Basic applications and programming for confirmatory factor analytic models*. Springer Science & Business Media.

Cai, L. (2012). flexMIRT: Flexible multilevel item factor analysis and test scoring [Computer software]. *Seattle, WA: Vector Psychometric Group, LLC*.

Campbell, J. M., Ruble, L. A., & Hammond, R. K. (2014). Comprehensive developmental assessment model. In L. A. Wilkinson (Ed.), *Autism spectrum disorder in children and adolescents: Evidence-based assessment and intervention in schools* (pp. 51–73). American Psychological Association. https://doi.org/10.1037/14338-004

Caron, C., & Rutter, M. (1991). Comorbidity in child psychopathology: Concepts, issues and research strategies. *Journal of Child Psychology and Psychiatry*, *32*(7), 1063-1080. https://doi.org/10.1111/j.1469-7610.1991.tb00350.x

Carruthers, S., Charman, T., El Hawi, N., Kim, Y. A., Randle, R., Lord, C., Pickles, A., & the PACT Consortium. (2021). Utility of the Autism Diagnostic Observation Schedule and the Brief Observation of Social and Communication Change for Measuring Outcomes for a Parent-Mediated Early Autism Intervention. *Autism Research*, *14*(2), 411-425. https://doi.org/10.1002/aur.2449

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, *1*(2), 245-276. https://doi.org/10.1207/s15327906mbr0102_10

Charman, T., & Gotham, K. (2013). Measurement Issues: Screening and diagnostic instruments for autism spectrum disorders – lessons from research and practice. *Child and Adolescent Mental Health*, *18*(1), 52-63. https://doi.org/10.1111/j.1475-3588.2012.00664.x

Chebli, S. S., Martin, V., & Lanovaz, M. J. (2016). Prevalence of stereotypy in individuals with developmental disabilities: A systematic review. *Review Journal of Autism and Developmental Disorders*, *3*(2), 107-118. https://doi.org/10.1007/s40489-016-0069-x

Child, D. (1990). *The essentials of factor analysis*. Cassell Educational.

Chowdhury, M., Benson, B. A., & Hillier, A. (2010). Changes in restricted repetitive behaviors with age: A study of high-functioning adults with autism spectrum disorders. *Research in Autism Spectrum Disorders*, *4*(2), 210-216. https://doi.org/10.1016/j.rasd.2009.09.006

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, *6*(4), 284-290. https://psycnet.apa.org/doi/10.1037/1040-3590.6.4.284

Clark, E., Radley, K., & Phosaly, L. (2014). Best practices in assessment and intervention of children with high functioning autism spectrum disorders. In P. L. Harrison & A. Thomas (Eds.), *Best practices in school psychology: Data-based and collaborative decision making* (pp. 417-431). NASP Publications.

Cohen, D. J., & Volkmar, F. R. (1997). *Handbook of Autism and Pervasive Developmental Disorders*. John Wiley & Sons Inc.

Cohen, I. L., Schmidt-Lackner, S., Romanczyk, R., & Sudhalter, V. (2003). The PDD Behavior Inventory: A rating scale for assessing response to intervention in children with pervasive developmental disorder. *Journal of Autism and Developmental Disorders*, *33*, 31-45. https://doi.org/10.1023/A:1022226403878

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences (2nd ed.). Routledge. https://doi.org/10.4324/9780203771587

Cole, D. A. (1987). Utility of confirmatory factor analysis in test validation research. *Journal of Consulting and Clinical Psychology*, 55(4), 584–594. https://doi.org/10.1037/0022-006X.55.4.584

Comrey, A. L., & Lee, H. B. (1992). Interpretation and application of factor analytic results. In A. L. Comrey & H. B. Lee (Eds.), *A first course in factor analysis* (pp. 754-761). Lawrence Eribuam Associates.

Conners, C. K. (1997). *Conners' rating scales revised*. Multi-Health Systems, Incorporated.

Constantino, J. N., & Gruber, C. P. (2012). *Social Responsiveness Scale, Second Edition (SRS-2).* Torrance, CA; Western Psychological Services.

Costello, A. B., & Osborne, J. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research, and Evaluation*, *10*(1), 7. https://doi.org/10.7275/jyj1-4868

Crapper, L., & Ernst, C. (2015). Comparative Analysis of Self-Injury in People with Psychopathology or Neurodevelopmental Disorders. *Pediatric Clinics of North America*, *62*(3), 619-631. https://doi.org/10.1016/j.pcl.2015.03.001

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297-334. https://doi.org/10.1007/BF02310555

Cuccaro, M. L., Brinkley, J., Abramson, R. K., Hall, A., Wright, H. H., Hussman, J. P., ... & Pericak-Vance, M. A. (2007). Autism in African American families: Clinical-phenotypic findings. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, *144*(8), 1022-1026. https://doi.org/10.1002/ajmg.b.30535

Cuccaro, M. L., Shao, Y., Grubber, J., Slifer, M., Wolpert, C. M., Donnelly, S. L., Abramson, R. K., Ravan, S. A., Wright, H. H., DeLong, G. R., & Pericak-Vance, M. A. (2003). Factor analysis of restricted and repetitive behaviors in autism using the Autism Diagnostic Interview-Revised. *Child Psychiatry and Human Development*, *34*(1), 3-17. https://doi.org/10.1023/A:1025321707947

Culp, R. E., Howell, C. S., Culp, A. M., & Blankemeyer, M. (2001). Maltreated children's emotional and behavioral problems: Do teachers and parents see the same things?. *Journal of Child and Family Studies*, *10*(1), 39-50. https://doi.org/10.1023/A:1016631111859

Delmolino, L., & Harris, S. L. (2012). Matching children on the autism spectrum to classrooms: A guide for parents and professionals. *Journal of Autism and Developmental Disorders*, *42*(6), 1197-1204. https://doi.org/10.1007/s10803-011-1298-6

Deprey, L., & Ozonoff, S. (2018). Assessment of comorbid psychiatric conditions in autism spectrum disorders. In S. Goldstein & S. Ozonoff (Eds), *Assessment of Autism Spectrum Disorder, Second Edition* (pp. 290-317). The Guilford Press.

Devine, D. P. (2014). Self-injurious behaviour in autistic children: a neuro-developmental theory of social and environmental isolation. *Psychopharmacology*, *231*(6), 979-997. https://doi.org/10.1007/s00213-013-3279-2

DiStefano, C., & Morgan, G. B. (2014). A comparison of diagonal weighted least squares robust estimation techniques for ordinal data. *Structural Equation Modeling: A Multidisciplinary Journal*, *21*(3), 425-438. https://doi.org/10.1080/10705511.2014.915373

Dworzynski, K., Happé, F., Bolton, P., & Ronald, A. (2009). Relationship between symptom domains in autism spectrum disorders: A population based twin study. *Journal of Autism and Developmental Disorders*, *39*(8), 1197–1210. https://doi.org/10.1007/s10803-009-0736-1

Esbensen, A. J., Seltzer, M. M., Lam, K. S. L., & Bodfish, J. W. (2009). Age-related differences in restricted repetitive behaviors in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *39*(1), 57-66. https://doi.org/10.1007/s10803-008-0599-x

Esler, A. N., & Ruble, L.A. (2015) DSM-5 Diagnostic criteria for autism spectrum disorder with implications for school psychologists, *International Journal of School and Educational Psychology*, *3*(1), 1-15. https://doi.org/10.1080/21683603.2014.890148

Evans, D. W., Leckman, J. F., Carter, A., Reznick, J. S., Henshaw, D., King, R. A., & Pauls, D. (1997). Ritual, habit, and perfectionism: The prevalence and development of compulsive-like behavior in normal young children. *Child Development*, 58-68. https://doi.org/10.2307/1131925

Fabrigar, L. R., & Wegener, D. T. (2012). *Exploratory factor analysis*. Oxford University Press. https://doi.org/10.1093/acprof:osobl/9780199734177.001.0001

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, *4*(3), 272- 299. https://doi.org/10.1037/1082-989x.4.3.272

Falkmer, T., Anderson, K., Falkmer, M., & Horlin, C. (2013). Diagnostic procedures in autism spectrum disorders: a systematic literature review. *European Child and Adolescent Psychiatry*, *22*(6), 329-340. https://doi.org/10.1007/s00787-013-0375-0

Feliciano, P., Daniels, A. M., Snyder, L. G., Beaumont, A., Camba, A., Esler, A., Gulsrud, A. G., Mason, A., Gutierrez, A., Nicholson, A., Paolicelli, A. M., McKenzie, A. P., Rachubinski, A. L., Stephens, A. N., Simon, A. R., Stedman, A., Shocklee, A. D., Swanson, A., Finucane, B., … Chung, W. K. (2018). SPARK: a US cohort of 50,000 families to accelerate autism research. *Neuron*, *97*(3), 488-493. https://doi.org/10.1016/j.neuron.2018.01.015

Fischbach, G. D., & Lord, C. (2010). The Simons Simplex Collection: a resource for identification of autism genetic risk factors. *Neuron*, *68*(2), 192-195. https://doi.org/10.1016/j.neuron.2010.10.006

Flessner, C. A., Francazio, S., Murphy, Y. E., & Brennan, E. (2015). An examination of executive functioning in young adults exhibiting body-focused repetitive behaviors. *The Journal of Nervous and Mental Disease*, *203*(7), 555-558. https://doi.org/10.1097/nmd.0000000000000327

Foldnes, N., & Olsson, U. H. (2015). Correcting too much or too little? The performance of three chi-square corrections. *Multivariate behavioral research*, *50*(5), 533-543. https://doi.org/10.1080/00273171.2015.1036964

Floyd, F. J., & Widman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. Psychological Assessment, *7*(3), 286 –299. https://doi.org/10.1037/1040-3590.7.3.286

Frazier, T. W., Georgiades, S., Bishop, S. L., & Hardan, A. Y. (2014). Behavioral and cognitive characteristics of females and males with autism in the Simons Simplex Collection. *Journal of the American Academy of Child & Adolescent Psychiatry*, *53*(3), 329-340.e3. https://doi.org/10.1016/j.jaac.2013.12.004

Frith, C. D., & Done, D. J. (1990). Stereotyped behaviour in madness and in health. In S. J. Cooper & C. T. Dourish (Eds.), *Neurobiology of Stereotyped Behaviour* (pp. 232–259). Clarendon Press/Oxford University Press.

Fulceri, F., Narzisi, A., Apicella, F., Balboni, G., Baldini, S., Brocchini, J., Domenici, I., Cerullo, S., Igliozzi, R., Cosenza, A., Tancredi, R., Muratori, F., & Calderoni, S. (2016). Application of the Repetitive Behavior Scale-Revised – Italian version in preschoolers with autism spectrum disorder. *Research in Developmental Disabilities*, *48*, 43-52. https://doi.org/10.1016/j.ridd.2015.10.015

Fung, L. K., Mahajan, R., Nozzolillo, A., Bernal, P., Krasner, A., Jo, B., Coury, D., Whitaker, A., Veenstra-Vanderweele, J., & Hardan, A. Y. (2016). Pharmacologic treatment of severe irritability and problem behaviors in autism: a systematic review and meta-analysis. *Pediatrics*, *137* (Supplement_2), S124-S135. https://doi.org/10.1542/peds.2015-2851K

Fyfield, R. (2014). *The Rise and Fall of Repetitive Behaviors in a Community Sample of Infants and Toddlers* (Doctoral dissertation, Cardiff University).

Gadermann, A. M., Guhn, M., & Zumbo, B. D. (2012). Estimating ordinal reliability for Likert-type and ordinal item response data: A conceptual, empirical, and practical guide. *Practical Assessment, Research, and Evaluation*, *17*(1), 3. https://doi.org/10.7275/n560-j767

Gadow, K., & Sprafkin, J. (2010). Child and adolescent symptom inventory 4R: Screening and norms manual. Checkmate Plus.

Gagnon, S. G., Nagle, R. J., & Nickerson, A. B. (2007). Parent and teacher ratings of peer interactive play and social-emotional development of preschool children at risk. *Journal of Early Intervention*, *29*(3), 228-242. https://doi.org/10.1177/105381510702900303

Gal, E., & Yirmiya, N. (2021). Introduction: Repetitive and Restricted Behaviors and Interests in Autism Spectrum Disorders. *Repetitive and Restricted Behaviors and Interests in Autism Spectrum Disorders,* 1-11. https://doi.org/10.1007/978-3-030-66445-9_1

Georgiades, S., Papageorgiou, V., & Anagnostou, E. (2010). Brief report: Repetitive behaviours in Greek individuals with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *40*(7), 903-906. https://doi.org/10.1007/s10803-009-0927-9

Georgiades, S., Szatmari, P., Zwaigenbaum, L., Duku, E., Bryson, S., Bryson, S., Roberts, W., Goldberg, J., & Mahoney, W. (2007). Structure of the autism symptom phenotype: A proposed multidimensional model. *Journal of the American Academy of Child and Adolescent Psychiatry*, *46*(2), 188-196. https://doi.org/10.1097/01.chi.0000242236.90763.7f

Gilliam, J. E. (2014). *Gilliam Autism Rating Scale–Third Edition (GARS-3)*. Pro-Ed.

Glorfeld, L. W. (1995). An improvement on Horn's parallel analysis methodology for selecting the correct number of factors to retain. *Educational and psychological measurement*, *55*(3), 377-393. https://doi.org/10.1177/0013164495055003002

Goldin, R. L., Matson, J. L., & Cervantes, P. E. (2014). The effect of intellectual disability on the presence of comorbid symptoms in children and adolescents with autism spectrum disorder. *Research in Autism Spectrum Disorders*, *8*(11), 1552-1556. https://doi.org/10.1016/j.rasd.2014.08.006

Goldstein, S. (2018). Historical perspective and overview. In S. Goldstein & S. Ozonoff (Eds), *Assessment of autism spectrum disorder, second edition* (pp. 1-25). The Guilford Press.

Goldstein, S., & Naglieri, J. A. (2009). *Autism Spectrum Rating Scales (ASRS)*. Multi-Health Systems. https://doi.org/10.1037/t04938-000

Gorsuch, R.L. (1983). Factor analysis (2nd Ed.). Lawrence Erlbaum Associates.

Gotham, K., Risi, S., Pickles, A., & Lord, C. (2007). The Autism Diagnostic Observation Schedule: revised algorithms for improved diagnostic validity. *Journal of Autism and Developmental Disorders*, *37*(4), 613-627. https://doi.org/10.1007/s10803-006-0280-1

Grant, J. E., & Stein, D. J. (2014). Body-focused repetitive behavior disorders in ICD-11. *Brazilian Journal of Psychiatry*, *36*, 59-64. https://doi.org/10.1590/1516-4446-2013-1228

Gross, D., Fogg, L., Garvey, C., & Julion, W. (2004). Behavior problems in young children: An analysis of cross-informant agreements and disagreements. *Research in Nursing and Health*, *27*(6), 413-425. https://doi.org/10.1002/nur.20040

Guillemin, F., Bombardier, C., & Beaton, D. (1993). Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *Journal of Clinical Epidemiology*, *46*(12), 1417-1432. https://doi.org/10.1016/0895-4356(93)90142-N

Guttman, L. (1954). Some necessary conditions for common-factor analysis. *Psychometrika*, *19*(2), 149-161. https://doi.org/10.1007/bf02289162

Hardy, K. K., Kollins, S. H., Murray, D. W., Riddle, M. A., Greenhill, L., Cunningham, C., Abikoff, H. B., McCracken, J. T., Vitiello, B., Davies, M., McGough, J. J., Posner, K., Skrobala, A. M., Swanson, J. M., Wigal, T., Wigal, S. B., Ghuman, J. K., & Chuang, S. Z. (2007). Factor structure of parent-and teacher-rated attention-deficit/hyperactivity disorder symptoms in the Preschoolers with Attention-Deficit/Hyperactivity Disorder Treatment Study (PATS). *Journal of Child and Adolescent Psychopharmacology*, *17*(5), 621-633. https://doi.org/10.1089/cap.2007.0073

Harrington, D. (2008). Creating a Confirmatory Factor Analysis Model. *Confirmatory Factor Analysis*, 21-35. https://doi.org/10.1093/acprof:oso/9780195339888.003.0002

Harrop, C., McBee, M., & Boyd, B. A. (2016). How are child restricted and repetitive behaviors associated with caregiver stress over time? A parallel process multilevel growth model. *Journal of Autism and Developmental Disorders*, *46*, 1773-1783. https://doi.org/10.1007/s10803-016-2707-7

Harstad, E. B., Fogler, J., Sideridis, G., Weas, S., Mauras, C., & Barbaresi, W. J. (2015). Comparing diagnostic outcomes of autism spectrum disorder using DSM-IV-TR and DSM-5 criteria. *Journal of Autism and Developmental Disorders*, *45*(5), 1437-1450. https://doi.org/10.1007/s10803-014-2306-4

Hatcher, L., & O'Rourke, N. (2013). *A step-by-step approach to using SAS for factor analysis and structural equation modeling*. SAS Institute.

He, H., Ye, N., Yi, L., & Yang, C. (2019). Validating the repetitive behavior scale-revised for children in China aged 3 to 8 with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *49*, 4941-4956. https://doi.org/10.1007/s10803-019-04210-x

Herlihy, L. E., Brooks, B., Dumont-Mathieu, T., Barton, M. L., Fein, D., Chen, C. M., & Robins, D. L. (2014). Standardized screening facilitates timely diagnosis of autism spectrum disorders in a diverse sample of low-risk toddlers. *Journal of Developmental and Behavioral Pediatrics: JDBP*, *35*(2), 85. https://doi.org/10.1097%2FDBP.0000000000000014

Honey, E., Rodgers, J., & McConachie, H. (2012). Measurement of restricted and repetitive behaviour in children with autism spectrum disorder: Selecting a questionnaire or interview. *Research in Autism Spectrum Disorders*, *6*(2), 757-776. https://doi.org/10.1016/j.rasd.2011.10.011

Hong, E., & Matson, J. L. (2021) Factors associated with restricted, repetitive behaviors and interests and diagnostic severity level ratings in young children with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 51(*12*),* 4644–4654. https://doi.org/10.1007/s10803-021-04905-0

Hooker, J. L., Dow, D., Morgan, L., Schatschneider, C., & Wetherby, A. M. (2019). Psychometric analysis of the repetitive behavior scale-revised using confirmatory factor analysis in children with autism. *Autism Research*, *12*(9), 1399-1410. https://doi.org/10.1002/aur.2159

Horlin, C., Falkmer, M., Parsons, R., Albrecht, M. A., & Falkmer, T. (2014). The cost of autism spectrum disorders. *PloS One*, *9*(9), e106552. https://doi.org/10.1371/journal.pone.0106552

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. Psychometrika, *30*(2), 179-185. https://doi.org/10.1007/BF02289447

Houghton, D. C., Alexander, J. R., Bauer, C. C., & Woods, D. W. (2018). Body-focused repetitive behaviors: More prevalent than once thought?. *Psychiatry Research*, *270*, 389-393. https://doi.org/10.1016/j.psychres.2018.10.002

Hoyt, W. T. (2000). Rater bias in psychological research: When is it a problem and what can we do about it? *Psychological Methods*, *5*(1), 64–86. https://doi.org/10.1037/1082-989X.5.1.64

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1-55. https://doi.org/10.1080/10705519909540118

Huerta, M., & Lord, C. (2012). Diagnostic evaluation of autism spectrum disorders. *Pediatric Clinics of North America*, *59*(1), 103-111. https://dx.doi.org/10.1016%2Fj.pcl.2011.10.018

Hughes, C., Oksanen, H., Taylor, A., Jackson, J., Murray, L., Caspi, A., & Moffitt, T. E. (2002). `I'm gonna beat you!' SNAP!: An observational paradigm for assessing young children's disruptive behaviour in competitive play. *Journal of Child Psychology and Psychiatry*, *43*(4), 507-516. https://doi.org/10.1111/1469-7610.00041

Hus, V., & Lord, C. (2014). The autism diagnostic observation schedule, module 4: revised algorithm and standardized severity scores. *Journal of Autism and Developmental Disorders*, *44*(8), 1996-2012. https://doi.org/10.1007/s10803-014-2080-3

IBM Corp. Released 2019. IBM SPSS Statistics for Windows, Version 26.0 IBM Corp.

Iacobucci, D. (2010). Structural equations modeling: Fit indices, sample size, and advanced topics. *Journal of Consumer Psychology*, *20*(1), 90-98. https://doi.org/10.1016/j.jcps.2009.09.003

Inada, N., Ito, H., Yasunaga, K., Kuroda, M., Iwanaga, R., Hagiwara, T., Tani, I., Yukihiro, R., Uchiyama, T., Ogasahara, K., Hara, K., Inoue, M., Murakami, T., Someki, F., Nakamura, K., Sugiyama, T., Uchida, H., Ichikawa, H., Kawakubo, Y., ... Tsujii, M. (2015). Psychometric properties of the RBS-R for individuals with autism spectrum disorder in Japan. *Research in Autism Spectrum Disorders*, *15-16*, 60-68. https://doi.org/10.1016/j.rasd.2015.01.002

Ingram, D. G., Takahashi, T. N., & Miles, J. H. (2008). Defining autism subgroups: A taxometric solution. *Journal of Autism and Developmental Disorders*, *38*(5), 950-960. https://doi.org/10.1007/s10803-007-0469-y

International Test Commission. (2017). *The ITC Guidelines for Translating and Adapting Tests* (2nd ed.). Retrieved from www.InTestCom.org

Jackson, D. L., Gillaspy, J. A., & Purc-Stephenson, R. (2009). Reporting practices in confirmatory factor analysis: an overview and some recommendations. *Psychological Methods*, *14*(1), 6-23. https://doi.org/10.1037/a0014694

Jennrich, R. I., & Sampson, P. F. (1966). Rotation for simple loadings. *Psychometrika*, *31*(3), 313-323. https://doi.org/10.1007/BF02289465

Jepsen, M. I., Gray, K. M., & Taffe, J. R. (2012). Agreement in multi-informant assessment of behaviour and emotional problems and social functioning in adolescents with Autistic and Asperger's Disorder. *Research in Autism Spectrum Disorders*, *6*(3), 1091-1098. https://doi.org/10.1016/j.rasd.2012.02.008

Johnson, M., McConachie, H., Watson, M., Freeston, M., & Le Couteur, A. S. (2006). Profile of repetitive behaviours in children with ASD from 2 to 9 years. Paper presented at the European Academy of Childhood Disability conference.

Joseph, L., Thurm, A., Farmer, C., & Shumway, S. (2013). Repetitive behavior and restricted interests in young children with autism: Comparisons with controls and stability over 2 years: Restricted/repetitive behavior over time. *Autism Research*, *6*(6), 584-595. https://doi.org/10.1002/aur.1316

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, *20*(1), 141-151. https://doi.org/10.1177/001316446002000116

Kaiser, H. F., & Rice, J. (1974). Little jiffy, mark IV. *Educational and Psychological Measurement*, *34*(1), 111-117. https://doi.org/10.1177/001316447403400115

Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child*, *2*(3), 217-250.

Karst, J. S., & Van Hecke, A. V. (2012). Parent and family impact of autism spectrum disorders: A review and proposed model for intervention evaluation. *Clinical Child and Family Psychology Review*, *15*(3), 247-277. https://doi.org/10.1007/s10567-012-0119-6

Kästel, I. S., Vllasaliu, L., Wellnitz, S., Cholemkery, H., Freitag, C. M., & Bast, N. (2020). Repetitive Behavior in Children and Adolescents: Psychometric Properties of the German Version of the Repetitive Behavior Scale-Revised. *Journal of Autism and Developmental Disorders*, *51*(4), 1224-1237. https://doi.org/10.1007/s10803-020-04588-z

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*(430), 773-795. https://doi.org/10.1080/01621459.1995.10476572

Kaat, A. J., & Lecavalier, L. (2013). Disruptive behavior disorders in children and adolescents with autism spectrum disorders: A review of the prevalence, presentation, and treatment. *Research in Autism Spectrum Disorders*, *7*(12), 1579-1594. https://doi.org/10.1016/j.rasd.2013.08.012

Kanakri, S. M., Shepley, M., Varni, J. W., & Tassinary, L. G. (2017). Noise and autism spectrum disorder in children: An exploratory survey. *Research in Developmental Disabilities, 63*, 85-94. https://doi.org/10.1016/j.ridd.2017.02.004

Kazdin, A. E. (2017). *Research Design in Clinical Psychology*. Cambridge University Press.

Kim, S. H., & Lord, C. (2010). Restricted and repetitive behaviors in toddlers and preschoolers with autism spectrum disorders based on the Autism Diagnostic Observation Schedule (ADOS). *Autism Research, 3*(4), 162–173. https://doi.org/10.1002/aur.142

Kim, S. H., & Lord, C. (2012). New autism diagnostic interview-revised algorithms for toddlers and young preschoolers from 12 to 47 months of age. *Journal of Autism and Developmental Disorders*, *42*(1), 82-93. https://doi.org/10.1007/s10803-011-1213-1

Klin, A., Danovitch, J. H., Merz, A. B., & Volkmar, F. R. (2007). Circumscribed interests in higher functioning individuals with autism spectrum disorders: An exploratory study. *Research and Practice for Persons with Severe Disabilities*, *32*(2), 89-100. https://doi.org/10.2511/rpsd.32.2.89

Kline, P. (1994). *An easy guide to factor analysis*. Routledge. https://doi.org/10.4324/9781315788135

Klopper, F., Testa, R., Pantelis, C., & Skafidas, E. (2017). A cluster analysis exploration of autism spectrum disorder subgroups in children without intellectual disability. *Research in Autism Spectrum Disorders*, *36*, 66-78. https://doi.org/10.1016/j.rasd.2017.01.006

Koegel, R. L., & Covert, A. (1972). The relationship of self-stimulation to learning in autistic children. *Journal of Applied Behavior Analysis*, 5(4), 381-387. https://doi.org/10.1901/jaba.1972.5-381

Koegel, R. L., Firestone, P. B., Kramme, K. W., & Dunlap, G. (1974). Increasing spontaneous play by suppressing self-stimulation in autistic children. *Journal of Applied Behavior Analysis,* 7(4), 521-528. https://doi.org/10.1901/jaba.1974.7-521

Kozlowski, A. M., Matson, J. L., Sipes, M., Hattier, M. A., & Bamburg, J. W. (2011). The relationship between psychopathology symptom clusters and the presence of comorbid psychopathology in individuals with severe to profound intellectual disability. *Research in Developmental Disabilities*, *32*(5), 1610-1614. https://doi.org/10.1016/j.ridd.2011.02.004

Krueger, R. F., & Markon, K. E. (2006). Reinterpreting comorbidity: A model-based approach to understanding and classifying psychopathology. *Annual Review of Clinical Psychology*, *2*(1), 111-133. https://doi.org/10.1146/annurev.clinpsy.2.022305.095213

Kurzius-Spencer, M., Pettygrove, S., Christensen, D., Pedersen, A. L., Cunniff, C., Meaney, F. J., Soke, G. N., Harrington, R. A., Durkin, M., & Rice, S. (2018). Behavioral problems in children with autism spectrum disorder with and without co-occurring intellectual disability. *Research in Autism Spectrum Disorders, 56*, 61-71. https://doi.org/10.1016/j.rasd.2018.09.002

Kuenssberg, R., & McKenzie, K. (2011). Confirmatory factor analysis of the Adult Asperger Assessment: The association of symptom domains within a clinical population. *Research in Developmental Disabilities, 32*(6), 2321–2329. https://doi.org/10.1016/j.ridd.2011.07.034

Lam, K. S. L. (2004). *The Repetitive Behavior Scale-Revised: Independent validation and the effects of subject variables* (Order No. 3148184). Available from ProQuest Dissertations & Theses Global. (305140064). https://proxy.lib.umich.edu/login?url=https://www.proquest.com/dissertations-theses/repetitive-behavior-scale-revised-independent/docview/305140064/se-2

Lam, K. S. L., & Aman, M. G. (2007). The Repetitive Behavior Scale-Revised: Independent validation in individuals with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *37*(5), 855-866. https://doi.org/10.1007/s10803-006-0213-z

Lam, K. S. L., Bodfish, J. W., & Piven, J. (2008). Evidence for three subtypes of repetitive behavior in autism that differ in familiarity and association with other symptoms. *Journal of Child Psychology and Psychiatry*, *49*(11), 1193-1200. https://doi.org/10.1111/j.1469-7610.2008.01944.x

Lane, B. R., Paynter, J., & Sharman, R. (2013). Parent and teacher ratings of adaptive and challenging behaviours in young children with autism spectrum disorders. *Research in Autism Spectrum Disorders*, *7*(10), 1196-1203. https://doi.org/10.1016/j.rasd.2013.07.011

Langen, M., Schnack, H. G., Nederveen, H., Bos, D., Lahuis, B. E., de Jonge, M. V., van Engeland, H., & Durston, S. (2009). Changes in the developmental trajectories of striatum in autism. *Biological Psychiatry, 66*(4), 327-333. https://doi.org/10.1016/j.neubiorev.2010.02.005

Lavelle, T. A., Weinstein, M. C., Newhouse, J. P., Munir, K., Kuhlthau, K. A., & Prosser, L. A. (2014). Economic burden of childhood autism spectrum disorders. *Pediatrics, 133*(3), e520-e529. https://doi.org/10.1542/peds.2013-0763

Le Couteur, A., Lord, C., & Rutter, M. (2003). *The Autism Diagnostic Interview—Revised (ADI-R).* Western Psychological Services.

Ledesma, R. D., & Valero-Mora, P. (2007). Determining the number of factors to retain in EFA: An easy-to-use computer program for carrying out parallel analysis. *Practical Assessment, Research, and Evaluation*, *12*(1), 2. https://doi.org/10.7275/wjnc-nm63

Leekam, S. R., Nieto, C., Libby, S. J., Wing, L., & Gould, J. (2007). Describing the sensory abnormalities of children and adults with autism. *Journal of Autism and Developmental Disorders, 37*(5), 894-910. https://doi.org/10.1007/s10803-006-0218-7

Leekam, S. R., Prior, M. R., & Uljarevic, M. (2011). Restricted and repetitive behaviors in autism spectrum disorders: a review of research in the last decade. *Psychological Bulletin*, *137*(4), 562-593. https://doi.org/10.1037/a0023341

Lewis, M. H., & Bodfish, J. W. (1998). Repetitive behavior disorders in autism. *Mental Retardation and Developmental Disabilities Research Reviews*, *4*(2), 80-89. https://doi.org/10.1002/(sici)1098-2779(1998)4:2<80::aid-mrdd4>3.0.co;2-0

Lewis, M. H., & Kim, S. J. (2009). The pathophysiology of restricted repetitive behavior. *Journal of Neurodevelopmental Disorders*, *1*(2), 114-132. https://doi.org/10.1007/s11689-009-9019-6

Loftin, R. L., Odom, S. L., & Lantz, J. F. (2008). Social interaction and repetitive motor behaviors. *Journal of Autism and Developmental Disorders*, *38*, 1124-1135. https://doi.org/10.1007/s10803-007-0499-5

Lord, C., & Bishop, S. L. (2015). Recent advances in autism research as reflected in DSM-5 criteria for autism spectrum disorder. *Annual Review of Clinical Psychology*, *11*(1), 53-70. https://doi.org/10.1146/annurev-clinpsy-032814-112745

Lord, C., Corsello, C., & Grzadzinski, R. (2014). Diagnostic instruments in autistic spectrum disorders. *Handbook of Autism and Pervasive Developmental Disorders* 4(2). https://doi.org/10.1002/9781118911389.hautc25

Lord, C., & Jones, R. M. (2012a). Annual Research Review: Re-thinking the classification of autism spectrum disorders. *Journal of Child Psychology and Psychiatry*, *53*(5), 490-509. https://doi.org/10.1111/j.1469-7610.2012.02547.x

Lord, C., Petkova, E., Hus, V., Gan, W., Lu, F., Martin, D. M., Ousley, O., Guy, L., Bernier, R., Gerdts, J., Algermissen, M., Whitaker, A., Sutcliffe, J. S., Warren, Z., Klin, A., Saulnier, C., Hanson, E., Hundley, R., Piggot, J., … Risi., S. (2012). A multisite study of the clinical diagnosis of different autism spectrum disorders. *Archives of General Psychiatry*, *69*(3), 306-313. https://dx.doi.org/10.1001/archgenpsychiatry.2011.148

Lord, C., Rutter, M., DiLavore, P. C., & Risi, S. (1999). *Autism Diagnostic Observation Schedule – WPS* (ADOS-WPS). Western Psychological Services.

Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., & Bishop, S. L. (2012b). *Autism Diagnostic Observation Schedule – Second Edition* (*ADOS-2*). Western Psychological Services.

Lord, C., Rutter, M., & Le Couteur, A. Autism Diagnostic Interview-Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, 24**,** 659–685 (1994). https://doi.org/10.1007/BF02172145

Lovaas, O. I. (1977). *The autistic child: Language development through behavior modification*. Irvington.

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, *4*(1), 84-89. https://doi.org/10.1037/1082-989X.4.1.84

Maenner, M. J., Shaw, K. A., Baio, J., Washington, A., Patrick, M., DiRienzo, M., Christensen, D. L., Wiggins, L. D., Pettygrove, S., Andrews, J. G., Lopez, M., Hudson, A., Baroud, T., Schwenk, Y., White, T., Rosenberg, C. R., Lee, L., Harrington, R. A., Hutson, M. Dietz, P.M. (2020). Prevalence of autism spectrum disorder among children aged 8 years — Autism and developmental disabilities monitoring network. *Morbidity and Mortality Weekly Report (MMWR) Surveillance Summaries*, *69*(4), 1–12. http://doi.org/10.15585/mmwr.ss6904a1

Maenner, M. J., Warren, Z., Williams, A. R., Amoakohene, E., Bakian, A. V., Bilder, D. A., Durkin, M. S., Fitzgerald, R. T., Furnier, S. M., Hughes, M. M., Ladd-Acosta, C. M., McArthur, D., Pas, E. T., Salinas, A., Vehorn, A., Williams, S., Esler, A., Grzybowski, A., Hall-Lande, J., … & Shaw, K. A. (2023). Prevalence and characteristics of autism spectrum disorder among children aged 8 years—Autism and developmental disabilities monitoring network, 11 sites, United States, 2020. *Morbidity and Mortality Weekly Report (MMWR) Surveillance Summaries*, *72*(2), 1-14. https://doi.org/10.15585%2Fmmwr.ss7202a1

Major, S. O., Seabra-Santos, M. J., & Martin, R. P. (2015). Are we talking about the same child? Parent-teacher ratings of preschoolers' social-emotional behaviors. *Psychology in the Schools*, *52*(8), 789-799. https://doi.org/10.1002/pits.21855

Malone, R. P., Delaney, M. A., Hyman, S. B., & Cater, J. R. (2007). Ziprasidone in adolescents with autism: an open-label pilot study. *Journal of Child and Adolescent Psychopharmacology*, *17*(6), 779-790. https://doi.org/10.1089/cap.2006.0126

Mandy, W., Charman, T., Gilmour, J., & Skuse, D. (2011). Toward specifying pervasive developmental disorder—not otherwise specified. *Autism Research*, *4*(2), 121-131. https://doi.org/10.1002/aur.178

Marsh, H. W., Guo, J., Dicke, T., Parker, P. D., & Craven, R. G. (2020). Confirmatory factor analysis (CFA), exploratory structural equation modeling (ESEM), and set-ESEM: Optimal balance between goodness of fit and parsimony. *Multivariate Behavioral Research*, *55*(1), 102-119. https://doi.org/10.1080/00273171.2019.1602503

Marsh, H. W., Muthén, B., Asparouhov, T., Lüdtke, O., Robitzsch, A., Morin, A. J., & Trautwein, U. (2009). Exploratory structural equation modeling, integrating CFA and EFA: Application to students' evaluations of university teaching. *Structural Equation Modeling: A Multidisciplinary Journal*, *16*(3), 439-476. https://doi.org/10.1080/10705510903008220

Martin, C., Pepa, L., & Lord, C. (2018). DSM-5 Diagnosis of autism spectrum disorder. In Goldstein, S. & Ozonoff, S. (Eds.), *Assessment of Autism Spectrum Disorder, Second Edition* (pp. 96 – 129). The Guilford Press.

Martínez-González, A. E., & Piqueras, J. A. (2017). Validation of the Repetitive Behavior Scale-Revised in Spanish-speakers participants with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *48*(1), 198-208. https://doi.org/10.1007/s10803-017-3276-0

Masi, A., DeMayo, M. M., Glozier, N., & Guastella, A. J. (2017) An overview of autism spectrum disorder, heterogeneity, and treatment options. *Neuroscience Bulletin, 33*. 183-193. https://doi.org/10.1007/s12264-017-0100-y

Matson, J. L., Dempsey, T., & Fodstad, J. C. (2009). Stereotypies and repetitive/restrictive behaviours in infants with autism and pervasive developmental disorder. *Developmental Neurorehabilitation*, *12*(3), 122–127. https://doi.org/10.1080/17518420902936730

Matson, J. L., & Goldin, R. L. (2013). Comorbidity and autism: Trends, topics and future directions. *Research in Autism Spectrum Disorders*, *7*(10), 1228-1233. https://doi.org/10.1016/j.rasd.2013.07.003

Matson, J. L., & Kozlowski, A. M. (2011). The increasing prevalence of autism spectrum disorders. *Research in Autism Spectrum Disorders*, *5*(1), 418-425. https://doi.org/10.1016/j.rasd.2010.06.004

Mayes, S. D., Calhoun, S. L., Murray, M. J., & Zahid, J. (2011). Variables associated with anxiety and depression in children with autism. *Journal of Developmental and Physical Disabilities*, *23*(4), 325-337. https://doi.org/10.1007/s10882-011-9231-7

Mayes, S. D., Calhoun, S. L., Murray, M. J., Morrow, J. D., Yurich, K. K., Mahr, F., Cothren, S., Bouder, J. N., & Petersen, C. (2009). Comparison of scores on the Checklist for Autism Spectrum Disorder, Childhood Autism Rating Scale, and Gilliam Asperger's Disorder Scale for children with low functioning autism, high functioning autism, Asperger's disorder, ADHD, and typical development. *Journal of Autism and Developmental Disorders*, *39*(12), 1682-1693. https://doi.org/10.1007/s10803-009-0812-6

McConachie, H., Parr, J. R., Glod, M., Hanratty, J., Livingstone, N., Oono, I. P., Robalino, S., Baird, G., Beresford, B., Charman, T., Garland, D., Green, J., Gringras, P., Jones, G., Law, J., LeCouteur, A. S., Macdonald, G., McColl, E. M., Morris, C….Wil liams, K. (2015). Systematic review of tools to measure outcomes for young children with autism spectrum disorder. *19*(41) https://doi.org/10.3310/hta19410

Mehling, M. H., & Tassé, M. J. (2016). Severity of autism spectrum disorders: Current conceptualization, and transition to DSM-5. *Journal of Autism and Developmental Disorders, 46*, 2000-2016. https://doi.org/10.1007/s10803-016-2731-7

Mirenda, P., Smith, I. M., Vaillancourt, T., Georgiades, S., Duku, E., Szatmari, P., & the Pathways in ASD Study Team. (2010). Validating the repetitive behavior scale-revised in young children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *40*(12), 1521-1530. https://doi.org/10.1007/s10803-010-1012-0

Mirwis, J. E. (2011). *Exploratory factor analysis of the Aberrant Behavior Checklist-Community (ABC-C) with a sample of individuals with autism spectrum disorders*. State University of New York at Buffalo.

Moon, T. K. (1996). The expectation-maximization algorithm. *IEEE Signal Processing Magazine*, 13(6), 47-60. https://doi.org/10.1109/79.543975

Mooney, E. L., Gray, K. M., Tonge, B. J., Sweeney, D. J., & Taffe, J. R. (2009). Factor analytic study of repetitive behaviours in young children with pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, *39*(5), 765-774. https://doi.org/10.1007/s10803-008-0680-5

Muliak, S. A. (1972). The fundamentals of factor analysis. McGraw-Hill.

Murphy, K. R., & Davidshofer, C. O. (1988). *Psychological testing: Principles and applications*. Prentice-Hall.

Murray, D. S., Ruble, L. A., Willis, H., & Molloy, C. A. (2009). Parent and teacher report of social skills in children with autism spectrum disorders.

Muthén L. K., & Muthén, B. O. (1998-2017). Mplus User's Guide. (7th ed.). Muthén & Muthén.

Murray, D. S., Ruble, L. A., Willis, H., & Molloy, C. A. (2009). Parent and teacher report of social skills in children with autism spectrum disorders. *Language, Speech, and Hearing Services in Schools*, *40*(2), 109–115. https://doi.org/10.1044/0161-1461(2008/07-0089)

Muthén, L. K., & Muthén, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling: A Multidisciplinary Journal*, *9*(4), 599-620. https://doi.org/10.1207/S15328007SEM0904_8

Myers, S. M., & Johnson, C. P., & the Council on Children with Disabilities. (2007). Management of children with autism spectrum disorders. *Pediatrics*, *120*(5), 1162-1182. https://doi.org/10.1542/peds.2007-2362

Nathanson, E. W., & Rispoli, K. M. (2022). School Psychologists' Assessment Practices for Students with Co-Occurring Anxiety and Autism Spectrum Disorder. *Journal of Applied School Psychology*, *38*(2), 177-204. https://doi.org/10.1080/15377903.2021.1941468

Newschaffer, C. J., Falb, M. D., & Gurney, J. G. (2005). National autism prevalence trends from United States special education data. *Pediatrics*, *115*(3). e277-e282. https://doi.org/10.1542/peds.2004-1958

Noland, R. M., & Gabriels, R. L. (2004). Screening and identifying children with autism spectrum disorders in the public school system: The development of a model process. *Journal of Autism and Developmental Disorders*, *34*(3), 265-277. https://doi.org/10.1023/B:JADD.0000029549.84385.44

Norris, M., & Lecavalier, L. (2010a). Screening accuracy of Level 2 autism spectrum disorder rating scales: A review of selected instruments. *Autism*, *14*, 263–284. https://doi.org/10.1177%2F1362361309348071

Norris, M., & Lecavalier, L. (2010). Evaluating the use of exploratory factor analysis in developmental disability psychological research. *Journal of Autism and Developmental Disorders*, *40*(1), 8-20. https://doi.org/10.1007/s10803-009-0816-2

Nunnally, J. (1978). *Psychometric theory* (2nd ed.). McGraw-Hill.

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

Nyatsanza, S., Shetty, T., Gregory, C., Lough, S., Dawson, K., & Hodges, J. R. (2003). A study of stereotypic behaviours in Alzheimer’s disease and frontal and temporal variant frontotemporal dementia. *Journal of Neurology, Neurosurgery & Psychiatry*, *74*(10), 1398-1402. https://doi.org/10.1136/jnnp.74.10.1398

O’Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer’s MAP test. *Behavior research methods, instruments, & computers*, *32*(3), 396-402. https://doi.org/10.3758/BF03200807

Osborne, J. W., & Banjanovic, E. S. (2016). *Exploratory factor analysis with SAS*. SAS Institute.

Ozonoff, S., Goodlin-Jones, B. L., & Solomon, M. (2005). Evidence-based assessment of autism spectrum disorders in children and adolescents. *Journal of Clinical Child and Adolescent Psychology*, *34*(3), 523-540. https://doi.org/10.1207/s15374424jccp3403_8

Papageorgiou, V., Georgiades, S., & Mavreas, V. (2008). Brief report: Cross-cultural evidence for the heterogeneity of the restricted, repetitive behaviours and interests domain of autism: A Greek study. *Journal of Autism and Developmental Disorders*, *38*(3), 558-561. https://doi.org/10.1007/s10803-007-0409-x

Pedhazur, E. J.., Schmelkin, L.P., 1991. Measurement, Design, and Analysis: An Integrated Approach. Lawrence Erlbaum Associates, Publishers.

Pett, M. A., Lackey, N. R., & Sullivan, J. J. (2003). *Making sense of factor analysis: The use of factor analysis for instrument development in health care research.* Sage Publications, Inc. https://doi.org/10.4135/9781412984898

Portney, L. G., & Watkins, M. P. (2009). *Foundations of clinical research: applications to practice* (Vol. 892). Pearson/Prentice Hall.

Posserud, M. B., Lundervold, A. J., & Gillberg, C. (2006). Autistic features in a total population of 7–9-year-old children assessed by the ASSQ (Autism Spectrum Screening Questionnaire). *Journal of Child Psychology and Psychiatry*, *47*(2), 167-175. https://doi.org/10.1111/j.1469-7610.2005.01462.x

Powell, S. B., Bodfish, J. W., Parker, D. E., Crawford, T. W., & Lewis, M. H. (1996). Self-restraint and self-injury: Occurrence and motivational significance. *American Journal on Mental Retardation, 101*, 41-48.

Preacher, K. J., & MacCallum, R. C. (2003). Repairing Tom Swift's electric factor analysis machine. *Understanding statistics: Statistical Issues in Psychology, Education, and the Social Sciences, 2*(1), 13-43. https://doi.org/10.1207/S15328031US0201_02

Prior, M., & Macmillan, M. B. (1973). Maintenance of sameness in children with Kanner's syndrome. *Journal of Autism and Childhood Schizophrenia, 3*(2), 154-167. https://doi.org/10.1007/BF01537990

R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing. URL https://www.R-project.org/

Raftery, A. E. (1995). Bayesian Model Selection in Social Research. *Sociological Methodology, 25*, 111–163. https://doi.org/10.2307/271063

Rapin, I., Steinberg, M., & Waterhouse, L. (1999). Consistency in the ratings of behaviors of communicatively impaired autistic and non-autistic preschool children. *European Child and Adolescent Psychiatry, 8*(3), 214-224. https://doi.org/10.1007/s007870050132

Raulston, T. J., & Machalicek, W. (2018). Early intervention for repetitive behavior in autism spectrum disorder: A conceptual model. *Journal of Developmental and Physical Disabilities, 30*(1), 89-109. https://doi.org/10.1007/s10882-017-9566-9

Raykov, T., & Marcoulides, G. A. (2008). *An introduction to applied multivariate analysis*. Routledge. https://doi.org/10.4324/9780203809532

Rescorla, L. A., Bochicchio, L., Achenbach, T. M., Ivanova, M. Y., Almqvist, F., Begovac, I., Bilenberg, N., Bird, H., Dobrean, A., Erol, N., Fombonne, E., Fonseca A., Frigerio, A., Find, D. S., Lambert, M., Leung, P. W., Liu, X., Markovic, I… Verhulst, F. C. (2014). Parent–teacher agreement on children's problems in 21 societies. *Journal of Clinical Child & Adolescent Psychology, 43*(4), 627-642. https://doi.org/10.1080/15374416.2014.900719

Richards, C., Oliver, C., Nelson, L., & Moss, J. (2012). Self-injurious behaviour in individuals with autism spectrum disorder and intellectual disability. *Journal of Intellectual Disability Research, 56*(5), 476-489. https://doi.org/10.1111/j.1365-2788.2012.01537.x

Richler, J., Huerta, M., Bishop, S. L., & Lord, C. (2010). Developmental trajectories of restricted and repetitive behaviors and interests in children with autism spectrum disorders. *Development and Psychopathology, 22*(1), 55-69. https://doi.org/10.1017/S0954579409990265

Rispoli, M., Camargo, S., Machalicek, W., Lang, R., & Sigafoos, J. (2014). Functional communication training in the treatment of problem behavior maintained by access to rituals: Functional communication training. *Journal of Applied Behavior Analysis*, *47*(3), 580-593. https://doi.org/10.1002/jaba.130

Roberts, S., O'Connor, K., Aardema, F., & Bélanger, C. (2015). The impact of emotions on body-Focused repetitive behaviors: Evidence from a non-treatment-seeking sample. *Journal of Behavior Therapy and Experimental Psychiatry*, *46*, 189-197. https://doi.org/10.1016/j.jbtep.2014.10.007

Roid, G. H., & Sampers, J. L. (2004). *Merrill-Palmer-revised: Scales of development.* Stoelting.

Rojahn, J., Matlock, S. T., & Tasse, M.J. (2000) The Stereotyped Behavior Scale: psychometric properties and norms. *Research in Developmental Disabilities* 21: 437–454.

Rojahn, J., Schroeder, S. R., Mayo-Ortega, L., Oyama-Ganiko, R., LeBlanc, J., Marquis, J., & Berke, E. (2013). Validity and reliability of the Behavior Problems Inventory, the Aberrant Behavior Checklist, and the Repetitive Behavior Scale - Revised among infants and toddlers at risk for intellectual or developmental disabilities: A multi-method assessment approach. *Research in Developmental Disabilities*, *34*(5), 1804-1814. http://dx.doi.org/10.1016/j.ridd.2013.02.024

Rojas, D. C., Peterson, E., Winterrowd, E., Reite, M. L., Rogers, S. J., & Tregellas, J. R. (2006). Regional gray matter volumetric changes in autism associated with social and repetitive behavior symptoms. *BMC Psychiatry*, *6*(1), 1-13. https://doi.org/10.1186/1471-244X-6-56

Ronald, A., Simonoff, E., Kuntsi, J., Asherson, P., & Plomin, R. (2008). Evidence for overlapping genetic influences on autistic and ADHD behaviours in a community twin sample. *Journal of Child Psychology and Psychiatry*, *49*(5), 535-542. https://doi.org/10.1111/j.1469-7610.2007.01857.x

Russell, K. M., Frost, K. M., & Ingersoll, B. (2019). The relationship between subtypes of repetitive behaviors and anxiety in children with autism spectrum disorder. *Research in Autism Spectrum Disorders*, *62*, 48-54. https://doi.org/10.1016/j.rasd.2019.03.006

Rutter, M. L. (2011). Progress in understanding autism: 2007–2010. *Journal of Autism and Developmental Disorders*, *41*(4), 395-404. https://doi.org/10.1007/s10803-011-1184-2

Rutter, M. L., Bailey, A., & Lord, C. (2003). Social Communication Questionnaire (SCQ). Western Psychological Services.

Rydzewska, E., Fleming, M., Mackay, D., Young-Southward, G., Blacher, J., Ross Bolourian, Y., Widaman, K., & Cooper, S. A. (2021). General health status of youth with autism with and without intellectual disabilities transitioning from special education, and its relationship to personal and family circumstances: longitudinal cohort study. *International Journal of Developmental Disabilities*, 1-9. https://doi.org/10.1080/20473869.2021.1966600

Ryland, H. K., Hysing, M., Posserud, M. B., Gillberg, C., & Lundervold, A. J. (2012). Autism spectrum symptoms in children with neurological disorders. *Child and Adolescent Psychiatry and Mental Health*, *6*(1), 1-9. https://doi.org/10.1186/1753-2000-6-34

Sabus, A., Feinstein, J., Romani, P., Goldson, E., & Blackmer, A. (2019). Management of self-injurious behaviors in children with neurodevelopmental disorders: A pharmacotherapy overview. *Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy*, *39*(6), 645-664. https://doi.org/10.1002/phar.2238

Salvia, J., & Ysseldyke, J., & Witmer, S. E. (2017). *Assessment in special and inclusive education* (12th ed.). Houghton-Mifflin.

SAS Institute Inc 2013. SAS/ACCESS® 9.4 Interface to ADABAS: Reference. SAS Institute Inc.

Scahill, L., Aman, M. G., Lecavalier, L., Halladay, A. K., Bishop, S. L., Bodfish, J. W., Grondhuis, S., Jones, N., Horrigan, J. P., Cook, E. H., Handen, B. L., King, B. H., Pearson, D. A., McCracken, J. T., Sullivan, K. A. … Dawson, G. (2015). Measuring repetitive behaviors as a treatment endpoint in youth with autism spectrum disorder. *Autism*, *19*(1), 38-52. https://doi.org/10.1177/1362361313510069

Scahill, L., & Lord, C. (2004). Subject selection and characterization in clinical trials in children with autism. *CNS Spectrums*, *9*(1), 22-32. https://doi.org/10.1017/s1092852900008336

Scahill, L., McDougle, C. J., Williams, S. K., Dimitropoulos, A., Aman, M. G., McCracken, J. T., Tierney, E., Arnold, L. E., Cronin, P., Grados, M., Ghuman, J., Koenig, K., Lam, K. S. L., McGough, J., Posey, D. J., Ritz, L., Swiezy, N. B. … & Vitiello, B. (2006). Children's Yale-Brown Obsessive-Compulsive Scale modified for pervasive developmental disorders. *Journal of the American Academy of Child & Adolescent Psychiatry*, *45*(9), 1114-1123. https://doi.org/10.1097/01.chi.0000220854.79144.e7

Schaaf, R. C., & Lane, A. E. (2015). Toward a best-practice protocol for assessment of sensory features in ASD. *Journal of Autism and Developmental Disorders*, *45*(5), 1380-1395. https://doi.org/10.1007/s10803-014-2299-z

Schertz, H. H., Odom, S. L., Baggett, K. M., & Sideris, J. H. (2016). Parent-reported repetitive behavior in toddlers on the autism spectrum. *Journal of Autism and Developmental Disorders*, *46*(10), 3308-3316. https://doi.org/10.1007/s10803-016-2870-x

Schmitt, T. A. (2011). Current methodological considerations in exploratory and confirmatory factor analysis. *Journal of Psychoeducational Assessment*, *29*(4), 304-321. https://doi.org/10.1177/0734282911406653

Schopler, E. (Ed.). (1995). *Parent survival manual: A guide to crisis resolution in autism and related developmental disorders*. Springer Science & Business Media.

Schopler, E., Mesibov, G. B., & Hearsey, K. (1995). Structured teaching in the TEACCH system. In *Learning and cognition in autism* (pp. 243-268). Springer. https://doi.org/10.1007/978-1-4899-1286-2_13

Schopler, E., Van Bourgondien, M. E., Wellman, G. J., & Love, S. R. (2010). *The Childhood Autism Rating Scale, Second Edition* (CARS2). Western Psychological Services.

Schultz, T. M., & Berkson, G. (1995). Definition of abnormal focused affections and exploration of their relation to abnormal stereotyped behaviors. *American Journal on Mental Retardation*. *99*(4), 376–390.

Schwartz, I. S., & Davis, C. A. (2014). Best practices in early identification and services for children with autism spectrum disorders. *Best practices in school psychology: Data-based and collaborative decision making*, 405-416.

Seidman, I. G., & Yirmiya, N. (2018). Assessment of social behavior in autism spectrum disorder. In S. Goldstein & S. Ozonoff (Eds.), *Assessment of autism spectrum disorder* (pp. 147–178). Guilford Press.

Shattuck, P. T., Seltzer, M. M., Greenberg, J. S., Orsmond, G. I., Bolt, D., Kring, S., Lounds, J., & Lord, C. (2007). Change in Autism Symptoms and Maladaptive Behaviors in Adolescents and Adults with an autism spectrum disorder. *Journal of Autism and Developmental Disorders, 37*(9), 1735-1747. https://doi.org/10.1007/s10803-006-0307-7

Shaw K.A., Maenner M.J., Baio J, Washington, A., Christensen, D. L., Wiggins, L. D., Pettygrove, S., Andrews, J., White, T., Rosenberg, C. R., Constantino, J. N., Fitzgerald, R. T., Zahorodny, W, Shenouda, J., Daniels, J. L., Salinas, A., Durkin, M. S., & Dietz, P. M. (2020). Early identification of autism spectrum disorder among children aged 4 years — early autism and developmental disabilities monitoring network, six sites, United States, 2016. *MMWR Surveillance Summaries*, *69*(3) 1-11. http://dx.doi.org/10.15585/mmwr.ss6903a1

Shulman, C., & Bing, O. (2021). Sex/Gender and Repetitive and Restrictive Behaviors in Autism Spectrum Disorder. *Repetitive and Restricted Behaviors and Interests in Autism Spectrum Disorders: From Neurobiology to Behavior*, 91-114. https://doi.org/10.1007/978-3-030-66445-9_7

Siddiqui, E. U., Naeem, S. S., Naqvi, H., & Ahmed, B. (2012). Prevalence of body-focused repetitive behaviors in three large medical colleges of Karachi: a cross-sectional study. *Bio Med Central Research Notes*, *5*(1), 1-6. https://doi.org/10.1186/1756-0500-5-614

Stapleton, C.D. (1997). Basic Concepts and Procedures of Confirmatory Factor Analysis. In Proceedings of the Annual Meeting of the Southwest Educational Research Association, Austin, TX, USA, 23–25 January 1997.

Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate behavioral research*, *25*(2), 173-180. https://doi.org/10.1207/s15327906mbr2502_4

Steiger, J. H. (2016). Notes on the Steiger–Lind (1980) handout. *Structural equation modeling: A Multidisciplinary Journal*, *23*(6), 777-781. https://doi.org/10.1080/10705511.2016.1217487

Steiger, J. H. & Lind, J.C. (1980). Statistically based tests for the number of common factors. In *the annual meeting of the Psychometric Society. Iowa City, IA. 1980*.

Stone, S. L., Speltz, M. L., Collett, B., & Werler, M. M. (2013). Socioeconomic factors in relation to discrepancy in parent versus teacher ratings of child behavior. *Journal of Psychopathology and Behavioral Assessment*, *35*(3), 314-320. https://doi.org/10.1007/s10862-013-9348-3

Stratis, E. A., & Lecavalier, L. (2013). Restricted and repetitive behaviors and psychiatric symptoms in youth with autism spectrum disorders. *Research in Autism Spectrum Disorders*, *7*(6), 757-766. https://doi.org/10.1016/j.rasd.2013.02.017

Sturm, A., Huang, S., & Kuhfeld, M. (2022). Advancing methodologies to improve RRB outcome measures in autism research: Evaluation of the RBS-R. *Psychological assessment*, *34*(1), 30–42. https://doi.org/10.1037/pas0001062

Sparrow, S. S., Cicchetti, D. V., & Balla, D. A. (2005) *Vineland Adaptive Behavior Scales-Second Edition (VABS-II).* American Guidance Service.

Spearman, C. (1904). "General Intelligence," Objectively Determined and Measured. *The American Journal of Psychology*, *15*(2), 201–292. https://doi.org/10.2307/1412107

Stevens, J. (1996). Applied Multivariate Statistics for the Social Sciences. Lawerence Erlbaum Associates.

Suhr, D. D. (2005). Principal component analysis vs. exploratory factor analysis. *SAS Users Group International (SUGI) 30 Proceedings*, *203*, 230. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.444.2964&rep=rep1&type=pdf

Sullivan, J. R., & Riccio, C. A. (2007). Diagnostic group differences in parent and teacher ratings on the BRIEF and Conners' Scales. *Journal of Attention Disorders, 11*(3), 398-406. https://doi.org/10.1177/1087054707299399

Symons, F. J., Sperry, L. A., Dropik, P. L., & Bodfish, J. W. (2005). The early development of stereotypy and self-injury: a review of research methods. *Journal of Intellectual Disability Research*, *49*(2), 144-158. https://doi.org/10.1111/j.1365-2788.2004.00632.x

Szatmari, P., Bartolucci, G., & Bremner, R. (1989). Asperger's syndrome and autism: Comparison of early history and outcome. *Developmental Medicine & Child Neurology, 31*(6), 709-720. https://doi.org/10.1111/j.1469-8749.1989.tb04066.x

Szatmari, P., Georgiades, S., Bryson, S., Zwaigenbaum, L., Roberts, W., Mahoney, W., Goldberg, J., & Tuff, L. (2006). Investigating the structure of the restricted, repetitive behaviours and interests domain of autism: Investigating the structure of the RRBI domain of autism. *Journal of Child Psychology and Psychiatry, 47*(6), 582-590. https://doi.org/10.1111/j.1469-7610.2005.01537.x

Teng, E. J., Woods, D. W., Marcks, B. A., & Twohig, M. P. (2004). Body-focused repetitive behaviors: The proximal and distal effects of affective variables on behavioral expression. *Journal of Psychopathology and Behavioral Assessment, 26*, 55-64. https://doi.org/10.1023/B:JOBA.0000007456.24198.e4

Thelen, E. (1979). Rhythmical stereotypies in normal human infants. *Animal Behaviour, 27*, 699-715. https://doi.org/10.1016/0003-3472(79)90006-x

Toomey, J. A. (2008). *Construct validity of the Adapted Skillstreaming Survey for communicative individuals with autism spectrum disorders*. State University of New York at Buffalo.

Tracy, J. I., de Leon, J., Qureshi, G., McCann, E. M., McGrory, A., & Josiassen, R. C. (1996). Repetitive behaviors in schizophrenia: a single disturbance or discrete symptoms?. *Schizophrenia Research, 20*(1-2), 221-229. https://doi.org/10.1016/0920-9964(95)00104-2

Travers, B. G., Powell, P. S., Klinger, L. G., & Klinger, M. R. (2013). Motor difficulties in autism spectrum disorder: linking symptom severity and postural stability. *Journal of Autism and Developmental Disorders, 43*, 1568-1583. https://doi.org/10.1007/s10803-012-1702-x

Tureck, K., Matson, J. L., Cervantes, P., & Konst, M. J. (2014). An examination of the relationship between autism spectrum disorder, intellectual functioning, and comorbid symptoms in children. *Research in Developmental Disabilities, 35*(7), 1766-1772. https://doi.org/10.1016/j.ridd.2014.02.013

Turner, M. (1999). Annotation: Repetitive behaviour in autism: A review of psychological research. *The Journal of Child Psychology and Psychiatry and Allied Disciplines, 40*(6), 839-849. https://doi.org/10.1111/1469-7610.00502

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika, 38*(1), 1-10. https://doi.org/10.1007/BF02291170

Tziner, A., Murphy, K. R., & Cleveland, J. N. (2005). Contextual and rater factors affecting rating behavior. *Group and Organization Management, 30*(1), 89-98. https://doi.org/10.1177/1059601104267920 van’t Westeinde, A., Cauvet, É., Toro, R., Kuja-Halkola, R., Neufeld, J., Mevel, K., & Bölte, S. (2020). Sex differences in brain structure: a twin study on restricted and repetitive behaviors in twin pairs with and without autism. *Molecular Autism*, *11*(1), 1-20. https://doi.org/10.1186/s13229-019-0309-x

Van Wijngaarden-Cremers, V., Patricia, J. M., van Eeten, E., Groen, W. B., Van Deurzen, P. A., Oosterling, I. J., & Van der Gaag, R. J. (2014). Gender and age differences in the core triad of impairments in autism spectrum disorders: a systematic review and meta-analysis. *Journal of Autism and Developmental Disorders*, *44*(3), 627-635. https://doi.org/10.1007/s10803-013-1913-9

Varni, J. W., Lovaas, O. I., Koegel, R. L., & Everett, N. L. (1979). An analysis of observational learning in autistic and normal children. *Journal of Abnormal Child Psychology*, *7*(1), 31-43. https://doi.org/10.1007/BF00924508

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, *41*(3), 321-327. https://doi.org/10.1007/bf02293557

Velicer, W. F., Eaton, C. A., & Fava, J. L. (2000). Construct explication through factor or component analysis: A review and evaluation of alternative procedures for determining the number of factors or components. *Problems and Solutions in Human Assessment*, 41-71. https://doi.org/10.1007/978-1-4615-4397-8_3

Volker, M. A., Lopata, C., Vujnovic, R. K., Smerbeck, A. M., Toomey, J. A., Rodgers, J. D., Schiavo, A., & Thomeer, M. L. (2010). Comparison of the Bender Gestalt-II and VMI-V in samples of typical children and children with high-functioning autism spectrum disorders. *Journal of Psychoeducational Assessment*, *28*(3), 187-200. https://doi.org/10.1177/0734282909348216

Volkmar, F. R., Booth, L. L., McPartland, J. C., & Wiesner, L. A. (2014). Clinical evaluation in multidisciplinary settings. In F. R. Volkmar, S. J. Rogers, R. Paul, & K. A. Pelphrey (Eds.), *Handbook of autism and pervasive developmental disorders: Assessment, interventions, and policy, Fourth Edition* (pp. 661–672). John Wiley & Sons. https://doi.org/10.1002/9781118911389.hautc26

Volkmar, F., Chawarska, K., & Klin, A. (2005). Autism in infancy and early childhood. *Annual. Review of Psychology*, *56*, 315-336. https://doi.org/10.1146/annurev.psych.56.091103.070159

Volkmar, F. R., & Reichow, B. (2013). Autism in DSM-5: progress and challenges. *Molecular Autism*, *4*(1), 1-6. https://doi.org/10.1186/2040-2392-4-13

Watkins, M. W. (2018). Exploratory factor analysis: A guide to best practice. *Journal of Black Psychology*, *44*(3), 219-246. https://doi.org/10.1177/0095798418771807

Watt, N., Wetherby, A. M., Barber, A., & Morgan, L. (2008). Repetitive and stereotyped behaviors in children with autism spectrum disorders in the second year of life. *Journal* *of Autism and Developmental Disorders*, *38*(8), 1518-1533. https://doi.org/10.1007/s10803-007-0532-8

Weitlauf, A. S., Gotham K. O., Vehorn A. C., Warren, Z. E. (2014). Brief report: DSM-5 "levels of support:" A comment on discrepant conceptualizations of severity in ASD. *Journal of Autism and Developmental Disorders. 44*(2), 471-476. https://doi.org/10.1007/s10803-013-1882-z

Wheaton, B., Muthen, B., Alwin, D. F., & Summers, G. F. (1977). Assessing reliability and stability in panel models. *Sociological Methodology*, 8, 84–136. https://doi.org/10.2307/270754

Winsler, A., & Wallace, G. L. (2002). Behavior problems and social skills in preschool children: Parent-teacher agreement and relations with classroom observations. *Early Education and Development*, *13*(1), 41-58. https://doi.org/10.1207/s15566935eed1301_3

Witwer, A. N., & Lecavalier, L. (2008). Examining the validity of autism spectrum disorder subtypes. *Journal of Autism and Developmental Disorders, 38*(9), 1611-1624. https://doi.org/10.1007/s10803-008-0541-2

Wolff, J. J., Boyd, B. A., & Elison, J. T. (2016). A quantitative measure of restricted and repetitive behaviors for early childhood. *Journal of Neurodevelopmental Disorders*, *8*(1), 1-10. https://doi.org/10.1186/s11689-016-9161-x

Yirmiya, N., & Charman, T. (2010). The prodrome of autism: Early behavioral and biological signs, regression, peri- and post-natal development and genetics. *Journal of Child Psychology and Psychiatry, 51*(4), 432–458. https://doi.org/10.1111/j.1469-7610.2010.02214.x

Young, R. L., Lim, A. (2021) The Measurement of Restricted and Repetitive Behaviors in Autism Spectrum Disorder. In: Gal E., Yirmiya N. (Eds) *Repetitive and restricted behaviors and interests in autism spectrum disorders.* Autism and Child Psychopathology Series. Springer, Cham. https://doi.org/10.1007/978-3-030-66445-9_8

Zablotsky, B., Black, L. I., Maenner, M. J., Schieve, L. A., Danielson, M. L., Bitsko, R. H., Blumberg, S. J., Kogan, M. D., & Boyle, C. A. (2019). Prevalence and trends of developmental disabilities among children in the United States: 2009-2017. *Pediatrics*, 144(4), https://doi.org/10.1542/peds.2019-0811

Zheng, L., Grove, R., & Eapen, V. (2019). Spectrum or subtypes? A latent profile analysis of restricted and repetitive behaviours in autism. *Research in Autism Spectrum Disorders*, *57*, 46-54. https://doi.org/10.1016/j.rasd.2018.10.003

APPENDIX A: Study One Polychoric Correlation Matrix Estimates

Table A 1. Study One Inter-item Polychoric Correlation Matrix (N = 234)

Item	1	2	3	4	5	6	7	8	9	10
1	-									
2	0.609	-								
3	0.630	0.509	-							
4	0.656	0.509	0.650	-						
5	0.411	0.467	0.597	0.610	-					
6	0.516	0.541	0.711	0.591	0.633	-				
7	0.275	0.283	0.303	0.257	0.474	0.486	-			
8	0.218	0.146	0.278	0.361	0.540	0.436	0.683	-		
9	0.210	0.178	0.212	0.378	0.476	0.428	0.802	0.862	-	
10	0.152	0.226	0.296	0.259	0.353	0.443	0.481	0.349	0.342	-
11	0.302	0.316	0.195	0.351	0.342	0.406	0.640	0.538	0.567	0.601
12	0.187	0.363	0.249	0.310	0.409	0.370	0.589	0.523	0.560	0.478
13	0.121	0.361	0.269	0.128	0.344	0.493	0.435	0.338	0.406	0.517
14	0.181	0.162	0.044	0.211	0.253	0.097	0.520	0.391	0.491	0.369
15	0.118	0.143	0.215	0.327	0.199	0.260	0.119	0.196	-0.009	-0.017
16	0.056	0.044	0.192	0.363	0.272	0.225	0.198	0.274	0.111	0.197
17	0.126	0.024	0.262	0.428	0.290	0.326	0.183	0.393	0.132	0.038
18	0.161	0.125	0.177	0.382	0.290	0.156	0.081	0.364	0.158	0.278
19	-0.006	0.136	0.274	0.424	0.246	0.265	-0.060	0.264	0.010	-0.213
20	-0.030	0.055	0.085	0.386	0.226	0.123	0.186	0.293	0.216	0.225
21	0.097	0.205	0.227	0.473	0.269	0.276	0.323	0.365	0.341	0.144
22	0.358	0.193	0.396	0.465	0.427	0.504	0.280	0.560	0.394	0.163
23	0.133	0.177	0.301	0.328	0.258	0.410	0.281	0.188	0.288	0.074
24	0.051	0.149	0.241	0.339	0.233	0.320	0.170	0.198	0.200	0.100
25	0.130	0.219	0.276	0.378	0.346	0.349	0.254	0.242	0.110	0.096
26	-0.044	0.179	0.114	0.240	0.237	0.260	0.303	0.277	0.120	0.173
27	0.114	0.111	0.127	0.331	0.212	0.290	0.161	0.174	0.164	0.012
28	0.137	0.046	0.088	0.337	0.123	0.273	0.176	0.147	0.178	0.021
29	-0.128	0.083	0.032	0.299	0.189	0.136	0.118	0.314	0.070	-0.046
30	-0.018	-0.074	0.254	0.217	0.188	0.231	0.258	0.200	0.125	0.141
31	0.198	0.281	0.275	0.424	0.321	0.448	0.278	0.404	0.310	0.314
32	0.071	0.140	0.191	0.326	0.390	0.321	0.256	0.425	0.134	0.024
33	0.182	0.101	0.200	0.223	0.177	0.208	0.208	0.179	0.093	0.146
34	0.091	0.108	0.130	0.184	0.246	0.193	0.198	0.220	0.135	0.003
35	-0.127	0.067	0.376	0.490	0.331	0.304	0.378	0.315	0.226	0.147
36	0.291	0.213	0.339	0.505	0.385	0.439	0.228	0.365	0.392	0.137
37	0.009	0.002	0.143	0.247	0.287	0.334	0.356	0.427	0.459	0.231
38	-0.027	-0.030	0.103	0.236	0.161	0.260	0.294	0.351	0.327	0.151
39	0.072	0.016	0.133	0.291	0.317	0.331	0.375	0.436	0.412	0.084
40	0.166	0.189	0.173	0.207	0.211	0.317	0.148	0.270	0.301	0.020
41	0.274	0.319	0.234	0.346	0.252	0.377	0.185	0.238	0.207	0.015
42	0.338	0.295	0.379	0.542	0.583	0.510	0.198	0.407	0.351	0.141
43	0.277	0.279	0.333	0.481	0.679	0.514	0.232	0.424	0.362	0.171

Table A1 (cont'd)

Item	11	12	13	14	15	16	17	18	19	20	21
12	0.623										
13	0.442	0.456	-								
14	0.573	0.636	0.356	-							
15	0.226	0.388	0.085	0.336	-						
16	0.320	0.394	-0.021	0.367	0.850	-					
17	0.088	0.279	0.048	0.259	0.534	0.583	-				
18	0.247	0.507	-0.157	0.282	0.657	0.755	0.646	-			
19	0.239	0.219	-0.231	0.015	0.677	0.638	0.666	0.674	-		
20	0.415	0.339	0.070	0.254	0.633	0.628	0.472	0.698	0.609	-	
21	0.279	0.504	0.203	0.305	0.651	0.702	0.541	0.679	0.542	0.477	-
22	0.439	0.371	0.376	0.124	0.357	0.442	0.462	0.562	0.456	0.313	0.401
23	0.264	0.327	0.133	0.316	0.521	0.520	0.360	0.445	0.531	0.550	0.490
24	0.228	0.318	0.352	0.274	0.562	0.505	0.404	0.504	0.464	0.509	0.527
25	0.378	0.360	0.331	0.235	0.662	0.594	0.453	0.571	0.648	0.565	0.523
26	0.255	0.381	0.107	0.304	0.625	0.582	0.478	0.590	0.682	0.553	0.578
27	0.209	0.227	-0.036	0.256	0.619	0.640	0.353	0.501	0.575	0.472	0.637
28	0.243	0.224	0.165	0.137	0.461	0.438	0.385	0.451	0.482	0.360	0.510
29	0.126	0.322	-0.152	0.049	0.778	0.748	0.516	0.634	0.714	0.514	0.695
30	0.183	0.129	0.127	0.377	0.381	0.445	0.438	0.219	0.429	0.266	0.305
31	0.549	0.242	0.181	0.186	0.272	0.337	0.060	0.285	0.304	0.391	0.358
32	0.146	0.393	0.145	0.182	0.548	0.517	0.383	0.576	0.759	0.559	0.675
33	0.286	0.319	-0.078	0.241	0.507	0.585	0.174	0.562	0.615	0.519	0.456
34	0.299	0.319	-0.138	0.145	0.464	0.487	0.625	0.625	0.548	0.543	0.472
35	0.123	0.312	0.263	0.415	0.699	0.673	0.482	0.569	0.615	0.548	0.671
36	0.336	0.316	0.260	0.310	0.412	0.441	0.299	0.482	0.440	0.420	0.443
37	0.358	0.282	0.102	0.347	0.317	0.366	0.219	0.345	0.438	0.412	0.397
38	0.276	0.405	0.204	0.330	0.549	0.507	0.433	0.606	0.639	0.478	0.546
39	0.175	0.394	0.103	0.209	0.514	0.537	0.505	0.624	0.638	0.534	0.631
40	0.310	0.326	-0.056	0.203	0.412	0.462	0.160	0.272	0.663	0.323	0.482
41	0.250	0.306	0.023	0.149	0.357	0.342	0.121	0.329	0.421	0.392	0.373
42	0.343	0.411	0.038	0.216	0.348	0.445	0.396	0.442	0.500	0.280	0.434
43	0.400	0.340	-0.002	0.233	0.264	0.417	0.234	0.390	0.523	0.331	0.413

Item	22	23	24	25	26	27	28	29	30	31	32
23	0.299										
24	0.250	0.718	-								
25	0.490	0.711	0.850	-							
26	0.392	0.633	0.763	0.875	-						
27	0.285	0.618	0.639	0.751	0.771	-					
28	0.364	0.444	0.478	0.557	0.564	0.688	-				
29	0.265	0.463	0.519	0.597	0.706	0.596	0.455	-			
30	0.198	0.481	0.494	0.506	0.566	0.497	0.307	0.381	-		
31	0.337	0.397	0.279	0.394	0.320	0.388	0.287	0.453	0.371	-	

Table A1 (cont'd)

32	0.503	0.521	0.522	0.590	0.654	0.540	0.412	0.682	0.523	0.435	-
33	0.211	0.493	0.334	0.447	0.547	0.579	0.409	0.581	0.407	0.460	0.607
34	0.392	0.408	0.438	0.548	0.522	0.425	0.353	0.544	0.475	0.151	0.538
35	0.433	0.697	0.735	0.750	0.846	0.734	0.501	0.671	0.747	0.423	0.713
36	0.390	0.434	0.385	0.554	0.444	0.461	0.394	0.534	0.348	0.519	0.350
37	0.202	0.407	0.357	0.455	0.509	0.488	0.270	0.439	0.631	0.667	0.421
38	0.358	0.490	0.565	0.612	0.737	0.581	0.443	0.667	0.641	0.545	0.669
39	0.433	0.444	0.450	0.527	0.659	0.607	0.494	0.707	0.460	0.551	0.739
40	0.169	0.445	0.374	0.417	0.465	0.589	0.419	0.516	0.272	0.367	0.387
41	0.320	0.409	0.204	0.297	0.314	0.504	0.364	0.403	0.236	0.385	0.460
42	0.368	0.346	0.365	0.422	0.305	0.337	0.288	0.299	0.140	0.289	0.273
43	0.389	0.394	0.319	0.365	0.254	0.347	0.304	0.303	0.134	0.260	0.381

Item	33	34	35	36	37	38	39	40	41	42
34	0.439									
35	0.578	0.443	-							
36	0.388	0.393	0.491	-						
37	0.413	0.357	0.385	0.505	-					
38	0.595	0.570	0.764	0.504	0.653	-				
39	0.645	0.595	0.658	0.534	0.589	0.898	-			
40	0.457	0.379	0.421	0.454	0.365	0.421	0.583	-		
41	0.394	0.450	0.383	0.557	0.370	0.345	0.506	0.746	-	
42	0.234	0.304	0.380	0.352	0.265	0.244	0.265	0.581	0.575	-
43	0.312	0.511	0.273	0.357	0.200	0.158	0.326	0.542	0.495	0.767

APPENDIX B: Factor Solution Structure Matrices

Table B 1. Pattern Matrix for Study One Two-Factor Solution

Item	Item Content	Factor I	Factor II
18	CHECKING (Repeatedly checks doors, windows, drawers, appliances, clocks, locks, etc.)	**0.91**	-0.25
29	Insists that things remain in the same place(s) (e.g., toys, supplies, furniture, pictures, etc.)	**0.86**	-0.06
21	REPEATING (Need to repeat routine events; In/outdoor, up/down from chair, clothing on/off)	**0.83**	-0.15
16	COMPLETENESS (Must have doors opened or closed; Takes all items out of a container or area)	**0.82**	-0.10
15	ARRANGING/ORDERING (arranges certain objects in a particular pattern or place; Need for things to be even or symmetrical)	**0.81**	-0.07
25	SELF-CARE - BATHROOM AND DRESSING (Insists on specific order of activities or tasks related to using the bathroom, to watching, showering, bathing or dressing; Arranges items in a certain way in the bathroom or insists that bathroom items not be moved; Insists on wearing certain clothing items)	**0.81**	-0.06
38	Insists on same routine, household, school, or work schedule everyday	**0.78**	0.04
26	TRAVEL/TRANSPORTATION (Insists on taking certain routes/paths; Must sit in specific location in vehicles; Insists that certain items be present during travel, e.g., toy or material; Insists on seeing or touching certain things or places during travel such as a sign or store)	**0.77**	0.01
24	SLEEPING/BEDTIME (Insists on certain pre-bedtime routines; Arranges items in room "just so" prior to bedtime; Insists that certain items be present with him/her during sleep; Insists that another person be present prior to or during sleep)	**0.76**	-0.06
19	COUNTING (Counts items or objects; Counts to a certain number or in a certain way)	**0.71**	0.04
33	Insists on sitting at the same place	**0.71**	-0.04
39	Insists that specific things take place at specific times	**0.69**	0.11
23	EATING/MEALTIME (Strongly prefers/insists on eating/drinking only certain things; Eats or drinks items in a set order; Insists that meal related items are arranged in a certain way)	**0.66**	-0.04
35	Insists on using a particular door	**0.65**	0.06

Table B1 (cont'd)

17	WASHING/CLEANING (Excessively cleans certain body parts; Picks at lint or loose threads)	<u>**0.57**</u>	-0.05
32	Insists on walking in a particular pattern (e.g., straight line)	<u>**0.55**</u>	0.17
34	Dislikes changes in appearance or behavior of the people around him/her	<u>**0.54**</u>	0.24
27	PLAY/LEISURE (Insists on certain play activities; Follows a rigid routine during play/leisure; Insists that certain items be present/available during play/leisure; Insists that other persons do certain things during play)	<u>**0.53**</u>	0.23
40	Fascination, preoccupation with one subject or activity (e.g., trains, computers, weather, dinosaurs)	<u>**0.52**</u>	0.18
41	Strongly attached to one specific object	<u>**0.51**</u>	**0.37**
28	COMMUNICATION/SOCIAL INTERACTIONS (Repeats same topic(s) during social interactions; Repetitive questioning; Insists on certain topics of conversation; Insists that others say certain things or respond in certain ways during interactions)	<u>**0.49**</u>	0.00
37	Resists changing activities; Difficulty with transitions	<u>**0.49**</u>	**0.35**
20	HOARDING/SAVING (Collects, hoards, or hides specific items)	<u>**0.48**</u>	0.20
30	Objects to visiting new places	<u>**0.44**</u>	0.28
36	Likes the same CD, tape, record, or piece of music played continually; Likes same move/video or part of movie/video	<u>**0.44**</u>	0.22
7	HITS SELF WITH BODY PART (Hits or slaps head, face, or other body area)	-0.11	<u>**0.83**</u>
8	HITS SELF AGAINST SURFACE OR OBJECT (hits or bangs head or other body part on table, floor, or other surface)	-0.08	<u>**0.82**</u>
12	RUBS OR SCRATCHES SELF (Rubs or scratches marks on arms, leg, face, or torso)	-0.01	<u>**0.78**</u>
5	OBJECT USAGE (spins or twirls objects, twiddles or slaps or throws objects, lets objects fall out of hands)	-0.01	<u>**0.72**</u>
4	LOCOMOTION (turns in circles, whirls, jumps, bounces)	-0.02	<u>**0.72**</u>
6	SENSORY (covers eyes, looks closely or gazes at hands or objects, covers ears, smells or sniffs items, rubs surfaces)	-0.02	<u>**0.72**</u>
9	HITS SELF WITH OBJECT (Hits or bangs head or other body area with objects)	0.12	<u>**0.69**</u>
2	HEAD (rolls head, nods head, turns head)	0.03	<u>**0.68**</u>
10	BITES SELF (Bites hand, wrist, arm, lips, or tongue	-0.03	<u>**0.66**</u>
13	INSERTS FINGER OR OBJECT (eye-poking, ear-poking)	-0.09	<u>**0.62**</u>

Table B1 (cont'd)

1	WHOLE BODY (Body rocking, body swaying)	0.03	**<u>0.61</u>**
3	HAND/FINGER (Flaps hands, wiggles or flicks fingers, claps hands, waves or shakes hand or arm)	0.02	**<u>0.61</u>**
11	PULLS (pulls hair or skin)	-0.01	**<u>0.60</u>**
14	SKIN PICKING (picks at skin on face, hands, arms, legs, or torso)	-0.03	**<u>0.60</u>**
31	Becomes upset if interrupted in what he/she is doing	**0.33**	**<u>0.54</u>**
42	Preoccupation with part(s) of object rather than the whole object (e.g., buttons on clothes, wheels on cars)	**0.40**	**<u>0.45</u>**
43	Fascination, preoccupation with movement/things that move (e.g., fans, clocks)	**0.39**	**<u>0.40</u>**
22	TOUCH/TAP (Need to touch, tap, or rub items, surfaces, or people)	0.26	**<u>0.37</u>**

Table B 2. Pattern Matrix for Study One Three-Factor Solution

Item	Item Content	Factor I	Factor II	Factor III
18	CHECKING	**0.91**	-0.20	0.00
29	Insists that things remain in the same place(s)	**0.87**	-0.09	0.13
15	ARRANGING/ORDERING	**0.85**	-0.19	0.21
16	COMPLETENESS	**0.83**	-0.14	0.13
21	REPEATING	**0.81**	-0.09	0.02
25	SELF-CARE - BATHROOM AND DRESSING	**0.78**	0.01	0.03
19	COUNTING	**0.70**	0.00	0.15
26	TRAVEL/TRANSPORTATION	**0.68**	0.23	-0.12
33	Insists on sitting at the same place	**0.66**	0.09	-0.04
38	Insists on same routine, household, school, or work schedule everyday	**0.66**	**0.31**	-0.15
17	WASHING/CLEANING	**0.65**	-0.31	**0.32**
24	SLEEPING/BEDTIME	**0.65**	0.26	-0.23
35	Insists on using a particular door	**0.65**	0.02	0.14
23	EATING/MEALTIME	**0.56**	0.24	-0.19
39	Insists that specific things take place at specific times	**0.56**	**0.42**	-0.18
34	Dislikes changes in appearance or behavior of the people around him/her	**0.55**	0.07	**0.30**
32	Insists on walking in a particular pattern	**0.50**	0.21	0.08
40	Fascination, preoccupation with one subject or activity	**0.43**	**0.35**	-0.05
20	HOARDING/SAVING	**0.41**	0.29	0.03
36	Likes the same CD, tape, record, or piece of music played continually; Likes same move/video or part of movie/video	**0.40**	0.23	0.11
30	Objects to visiting new places	**0.38**	**0.31**	0.11
2	HEAD	-0.09	**0.73**	0.14
5	OBJECT USAGE	-0.11	**0.71**	0.21
1	WHOLE BODY	-0.09	**0.70**	0.09
4	LOCOMOTION	-0.09	**0.64**	0.27
43	Fascination, preoccupation with movement/things that move	0.25	**0.62**	-0.05
6	SENSORY	-0.09	**0.62**	0.30

Table B2 (cont'd)

42	Preoccupation with part(s) of object rather than the whole object (e.g., buttons on clothes, wheels on cars)	0.27	**<u>0.61</u>**	0.03
41	Strongly attached to one specific object	**0.37**	**<u>0.59</u>**	-0.04
3	HAND/FINGER	-0.06	**<u>0.57</u>**	0.20
8	HITS SELF AGAINST SURFACE OR OBJECT	-0.11	**<u>0.56</u>**	**0.47**
27	PLAY/LEISURE	**0.38**	**<u>0.54</u>**	-0.17
31	Becomes upset if interrupted in what he/she is doing	0.27	**<u>0.48</u>**	0.25
10	BITES SELF	-0.05	**<u>0.45</u>**	**0.39**
37	Resists changing activities; Difficulty with transitions	**0.41**	**<u>0.43</u>**	0.09
28	COMMUNICATION/SOCIAL INTERACTIONS	**0.35**	**<u>0.39</u>**	-0.32
22	TOUCH/TAP	0.23	<u>0.28</u>	0.22
14	SKIN PICKING	0.10	-0.04	**<u>0.79</u>**
12	RUBS OR SCRATCHES SELF	0.06	0.25	**<u>0.74</u>**
11	PULLS	0.09	0.06	**<u>0.70</u>**
13	INSERTS FINGER OR OBJECT	0.00	0.10	**<u>0.67</u>**
7	HITS SELF WITH BODY PART	-0.08	**0.40**	**<u>0.63</u>**
9	HITS SELF WITH OBJECT	0.13	**0.37**	**<u>0.53</u>**

Table B 3. Pattern Matrix for Study One Four-Factor Solution

Item	Item Content	Factor I	Factor II	Factor III	Factor IV
39	Insists that specific things take place at specific times	**0.91**	-0.08	-0.02	-0.03
38	Insists on same routine, household, school, or work schedule everyday	**0.90**	0.02	-0.09	-0.01
26	TRAVEL/TRANSPORTATION	**0.69**	0.21	0.01	-0.06
27	PLAY/LEISURE	**0.67**	-0.05	0.25	-0.09
33	Insists on sitting at the same place	**0.67**	0.17	-0.18	0.06
37	Resists changing activities; Difficulty with transitions	**0.65**	-0.04	0.09	0.21
30	Objects to visiting new places	**0.60**	-0.05	-0.02	0.23
41	Strongly attached to one specific object	**0.59**	0.01	**0.35**	0.01
24	SLEEPING/BEDTIME	**0.55**	**0.31**	0.18	-0.25
29	Insists that things remain in the same place(s)	**0.55**	**0.47**	-0.23	0.16
28	COMMUNICATION/SOCIAL INTERACTIONS	**0.51**	0.04	0.26	-0.31
32	Insists on walking in a particular pattern	**0.50**	0.16	0.02	0.14
35	Insists on using a particular door	**0.50**	0.29	-0.16	0.20
31	Becomes upset if interrupted in what he/she is doing	**0.49**	0.05	0.19	**0.35**
34	Dislikes changes in appearance or behavior of the people around him/her	**0.49**	0.19	-0.17	**0.40**
20	HOARDING/SAVING	**0.48**	0.09	0.09	0.09
40	Fascination, preoccupation with one subject or activity	**0.43**	0.17	0.24	-0.05
23	EATING/MEALTIME	**0.40**	**0.34**	0.24	-0.25
17	WASHING/CLEANING	-0.27	**0.89**	0.10	0.10
16	COMPLETENESS	0.14	**0.78**	0.07	-0.02
15	ARRANGING/ORDERING	0.24	**0.70**	-0.09	0.14
21	REPEATING	0.23	**0.69**	0.07	-0.10
19	COUNTING	0.12	**0.67**	0.18	0.02
18	CHECKING	**0.41**	**0.63**	-0.18	-0.04
25	SELF-CARE - BATHROOM AND DRESSING	**0.42**	**0.52**	0.01	-0.01

Table B3 (cont'd)

36	Likes the same CD, tape, record, or piece of music played continually; Likes same move/video or part of movie/video	0.24	<u>0.27</u>	0.21	0.07
4	LOCOMOTION	-0.11	0.09	<u>**0.74**</u>	0.17
1	WHOLE BODY	0.03	-0.01	<u>**0.73**</u>	0.01
2	HEAD	0.05	-0.02	<u>**0.73**</u>	0.07
5	OBJECT USAGE	0.01	-0.02	<u>**0.72**</u>	0.14
3	HAND/FINGER	-0.06	0.08	<u>**0.65**</u>	0.11
6	SENSORY	0.01	-0.02	<u>**0.60**</u>	0.26
43	Fascination, preoccupation with movement/things that move	**0.36**	0.08	<u>**0.54**</u>	-0.08
42	Preoccupation with part(s) of object rather than the whole object	**0.38**	0.08	<u>**0.49**</u>	0.01
22	TOUCH/TAP	0.01	0.29	<u>**0.37**</u>	0.14
11	PULLS	0.02	0.06	-0.09	<u>**0.78**</u>
12	RUBS OR SCRATCHES SELF	-0.10	0.16	0.21	<u>**0.75**</u>
14	SKIN PICKING	-0.30	**0.33**	0.07	<u>**0.75**</u>
7	HITS SELF WITH BODY PART	0.10	-0.13	0.18	<u>**0.72**</u>
13	INSERTS FINGER OR OBJECT	-0.14	0.10	0.06	<u>**0.69**</u>
9	HITS SELF WITH OBJECT	0.28	-0.06	0.10	<u>**0.63**</u>
8	HITS SELF AGAINST SURFACE OR OBJECT	0.27	-0.28	0.24	<u>**0.60**</u>
10	BITES SELF (Bites hand, wrist, arm, lips, or tongue	0.20	-0.18	0.21	<u>**0.48**</u>

Table B 4. Matrix for Study One Six-Factor Solution

Item	Item Content	Factor I	Factor II	Factor III	Factor IV	Factor V	Factor VI
39	Insists that specific things take place at specific times	**0.98**	-0.12	0.03	0.05	-0.10	0.03
38	Insists on same routine, household, school, or work schedule everyday	**0.90**	0.05	-0.09	-0.02	0.07	-0.07
37	Resists changing activities; Difficulty with transitions	**0.67**	-0.05	0.10	0.16	0.12	-0.01
33	Insists on sitting at the same place	**0.64**	0.15	-0.14	0.12	-0.09	0.20
30	Objects to visiting new places	**0.60**	-0.06	0.00	0.18	0.12	0.06
26	TRAVEL/TRANSPORTATION	**0.57**	0.28	-0.03	-0.11	0.11	0.04
30	Objects to visiting new places	**0.54**	0.09	-0.10	**0.46**	-0.09	0.25
28	COMMUNICATION/SOCIAL INTERACTIONS	**0.50**	-0.02	0.29	-0.15	-0.29	0.23
24	SLEEPING/BEDTIME	**0.49**	**0.41**	0.14	-0.24	0.04	-0.19
27	PLAY/LEISURE	**0.48**	0.03	0.20	-0.22	0.22	0.11
32	Insists on walking in a particular pattern	**0.46**	0.11	0.05	0.18	-0.07	0.25
23	EATING/MEALTIME	**0.42**	**0.41**	0.22	-0.19	-0.04	-0.29
31	Becomes upset if interrupted in what he/she is doing	**0.42**	-0.03	0.17	0.18	**0.31**	0.05
41	Strongly attached to one specific object	**0.39**	0.03	**0.31**	-0.09	0.14	**0.32**
17	WASHING/CLEANING	-0.17	**0.88**	0.15	0.28	-0.24	-0.17
16	COMPLETENESS	-0.04	**0.86**	0.03	-0.01	0.01	0.10
15	ARRANGING/ORDERING	0.02	**0.82**	-0.14	0.04	0.18	0.08
19	COUNTING	-0.06	**0.73**	0.15	0.02	-0.01	0.20
18	CHECKING	0.16	**0.72**	-0.22	-0.06	0.04	0.24
21	REPEATING	0.16	**0.70**	0.08	0.04	-0.23	0.14
25	SELF-CARE - BATHROOM AND DRESSING	**0.41**	**0.60**	-0.01	0.00	0.06	-0.26

Table B4 (cont'd)

29	Insists that things remain in the same place(s)	**0.37**	**<u>0.54</u>**	-0.25	0.09	0.13	0.15
2	HEAD	0.19	-0.11	**<u>0.79</u>**	0.15	-0.11	-0.02
1	WHOLE BODY	0.04	-0.06	**<u>0.75</u>**	0.04	-0.06	0.13
4	LOCOMOTION	-0.09	0.07	**<u>0.74</u>**	0.11	0.10	-0.03
5	OBJECT USAGE	-0.08	0.00	**<u>0.67</u>**	0.00	0.23	0.04
3	HAND/FINGER	-0.05	0.05	**<u>0.65</u>**	0.09	0.05	0.03
6	SENSORY	0.13	-0.09	**<u>0.64</u>**	0.27	0.02	-0.04
42	Preoccupation with part(s) of object rather than the whole object	0.13	0.12	**<u>0.44</u>**	-0.11	0.17	**0.38**
43	Fascination, preoccupation with movement/ things that move	0.07	0.14	**<u>0.46</u>**	-0.23	0.19	**0.37**
22	TOUCH/TAP	-0.08	0.29	**<u>0.36</u>**	0.10	0.05	0.19
13	INSERTS FINGER OR OBJECT	0.02	-0.05	0.15	**<u>0.70</u>**	0.00	0.13
14	SKIN PICKING	-0.17	0.25	0.12	**<u>0.69</u>**	0.14	-0.04
11	PULLS	0.16	-0.01	-0.03	**<u>0.67</u>**	0.23	-0.08
12	RUBS OR SCRATCHES SELF	-0.02	0.09	0.26	**<u>0.63</u>**	0.22	0.01
9	HITS SELF WITH OBJECT	0.01	0.11	-0.02	0.15	**<u>0.83</u>**	-0.04
8	HITS SELF AGAINST SURFACE OR OBJECT	0.00	-0.14	0.13	0.11	**<u>0.82</u>**	0.05
7	HITS SELF WITH BODY PART	-0.04	-0.03	0.10	0.30	**<u>0.75</u>**	-0.12
10	BITES SELF	0.10	-0.15	0.18	0.22	**<u>0.43</u>**	0.07
20	HOARDING/SAVING	0.25	0.18	0.02	-0.10	**<u>0.31</u>**	0.16
40	Fascination, preoccupation with one subject or activity	0.24	0.15	0.23	-0.05	-0.05	**<u>0.46</u>**
35	Insists on using a particular door	**0.33**	0.27	-0.15	0.18	0.01	**<u>0.44</u>**
36	Likes the same CD, tape, record, or piece of music played continually; Likes same move/video or part of movie/video	0.17	0.22	0.24	0.13	-0.12	**<u>0.31</u>**

Table B 5. Pattern Matrix for Study One Seven-Factor Solution

	Item Content	I	II	III	IV	V	VI	VII
16	COMPLETENESS	**0.90**	0.04	0.13	-0.09	-0.13	0.00	0.08
15	ARRANGING/ ORDERING	**0.84**	0.04	-0.09	0.00	-0.03	0.18	0.09
17	WASHING/ CLEANING	**0.84**	-0.20	0.11	**0.31**	-0.10	-0.24	0.28
18	CHECKING	**0.77**	0.21	-0.14	-0.13	0.04	0.03	-0.03
19	COUNTING	**0.70**	-0.19	0.07	0.09	**0.31**	-0.01	0.07
21	REPEATING	**0.70**	0.12	0.07	0.03	0.12	-0.23	0.11
29	Insists that things remain in the same place(s)	**0.54**	**0.30**	-0.27	0.10	0.17	0.13	0.07
25	SELF-CARE - BATHROOM AND DRESSING	**0.55**	0.23	-0.04	0.03	-0.04	0.06	**0.48**
39	Insists that specific things take place at specific times	-0.11	**0.92**	0.08	-0.03	0.03	-0.11	0.19
38	Insists on same routine, household, school, or work schedule everyday	0.05	**0.75**	-0.08	-0.05	0.06	0.06	**0.31**
33	Insists on sitting at the same place	0.20	**0.70**	-0.06	0.01	0.00	-0.10	-0.03
30	Objects to visiting new places	-0.02	**0.66**	0.08	0.07	-0.09	0.11	0.03
37	Resists changing activities; Difficulty with transitions	-0.04	**0.62**	0.12	0.10	0.02	0.12	0.16
34	Dislikes changes in appearance or behavior of the people around him/her	0.13	**0.61**	-0.08	**0.40**	0.10	-0.09	-0.14
32	Insists on walking in a particular pattern	0.17	**0.58**	0.14	0.06	0.00	-0.08	-0.11
31	Becomes upset if interrupted in what he/she is doing	0.02	**0.51**	0.28	0.05	-0.14	0.30	0.01
35	Insists on using a particular door	**0.35**	**0.50**	-0.05	0.06	0.08	0.00	-0.30
26	TRAVEL/ TRANSPORTATION	0.25	**0.34**	-0.10	-0.05	0.28	0.11	0.26

Table B5 (cont'd)

1	WHOLE BODY	-0.06	0.04	**<u>0.75</u>**	0.00	0.09	-0.07	0.01
2	HEAD	-0.13	0.12	**<u>0.75</u>**	0.13	0.06	-0.12	0.15
3	HAND/FINGER	0.08	5.00	**<u>0.74</u>**	-0.01	-0.16	0.04	0.02
4	LOCOMOTION	0.06	-0.11	**<u>0.73</u>**	0.09	-0.01	0.10	0.11
5	OBJECT USAGE	0.00	-0.09	**<u>0.69</u>**	-0.03	0.05	0.23	0.03
6	SENSORY	-0.13	0.04	**<u>0.56</u>**	0.29	0.12	0.02	0.14
43	Fascination, preoccupation with movement/ things that move	0.18	0.04	**<u>0.49</u>**	-0.27	0.27	0.18	-0.12
22	TOUCH/TAP	**0.33**	0.04	**<u>0.45</u>**	0.01	-0.06	0.04	-0.10
42	Preoccupation with part(s) of object rather than the whole object	0.15	0.12	**<u>0.46</u>**	-0.15	0.27	0.16	-0.13
11	PULLS	-0.05	0.13	-0.14	**<u>0.73</u>**	0.06	0.24	0.02
12	RUBS OR SCRATCHES SELF	0.07	0.00	0.18	**<u>0.65</u>**	0.05	0.22	-0.04
13	INSERTS FINGER OR OBJECT	-0.05	0.10	0.08	**<u>0.71</u>**	0.10	0.01	-0.19
14	SKIN PICKING	0.24	-0.06	0.11	**<u>0.68</u>**	-0.13	0.14	-0.07
40	Fascination, preoccupation with one subject or activity	0.09	-0.07	-0.02	0.13	**<u>0.84</u>**	-0.05	-0.07
28	COMMUNICATION/SOCIAL INTERACTIONS	-0.07	0.21	0.12	-0.04	**<u>0.57</u>**	-0.29	0.12
41	Strongly attached to one specific object	0.00	0.15	0.18	0.00	**<u>0.56</u>**	0.14	0.02
27	PLAY/LEISURE	-0.03	0.15	0.04	-0.10	**<u>0.51</u>**	0.22	0.22
20	HOARDING/ SAVING	0.13	-0.02	-0.13	0.03	**<u>0.48</u>**	**0.31**	0.10
36	Likes the same CD, tape, record, or piece of music played continually; Likes same move/video or part of movie/video	0.20	0.03	0.11	0.21	**<u>0.46</u>**	-0.12	-0.06
9	HITS SELF WITH OBJECT	0.13	0.05	0.05	0.10	-0.13	**<u>0.83</u>**	0.02
8	HITS SELF AGAINST SURFACE OR OBJECT	-0.15	-0.07	0.10	0.14	0.13	**<u>0.82</u>**	-0.03
7	HITS SELF WITH BODY PART	-0.05	-0.08	0.08	**0.32**	-0.03	**<u>0.75</u>**	0.07
10	BITES SELF	-0.12	0.19	0.24	0.15	-0.07	**<u>0.43</u>**	-0.10

Table B5 (cont'd)

23	EATING/ MEALTIME	**0.34**	0.15	0.15	-0.12	0.05	-0.04	**<u>0.56</u>**
24	SLEEPING/ BEDTIME	**0.35**	0.22	0.08	-0.18	0.12	0.04	**<u>0.49</u>**

APPENDIX C: Existing EFA Factor Models in the Current Literature

Figure C 1. Bodfish et al. (2000) Six-Factor Model Path Diagram

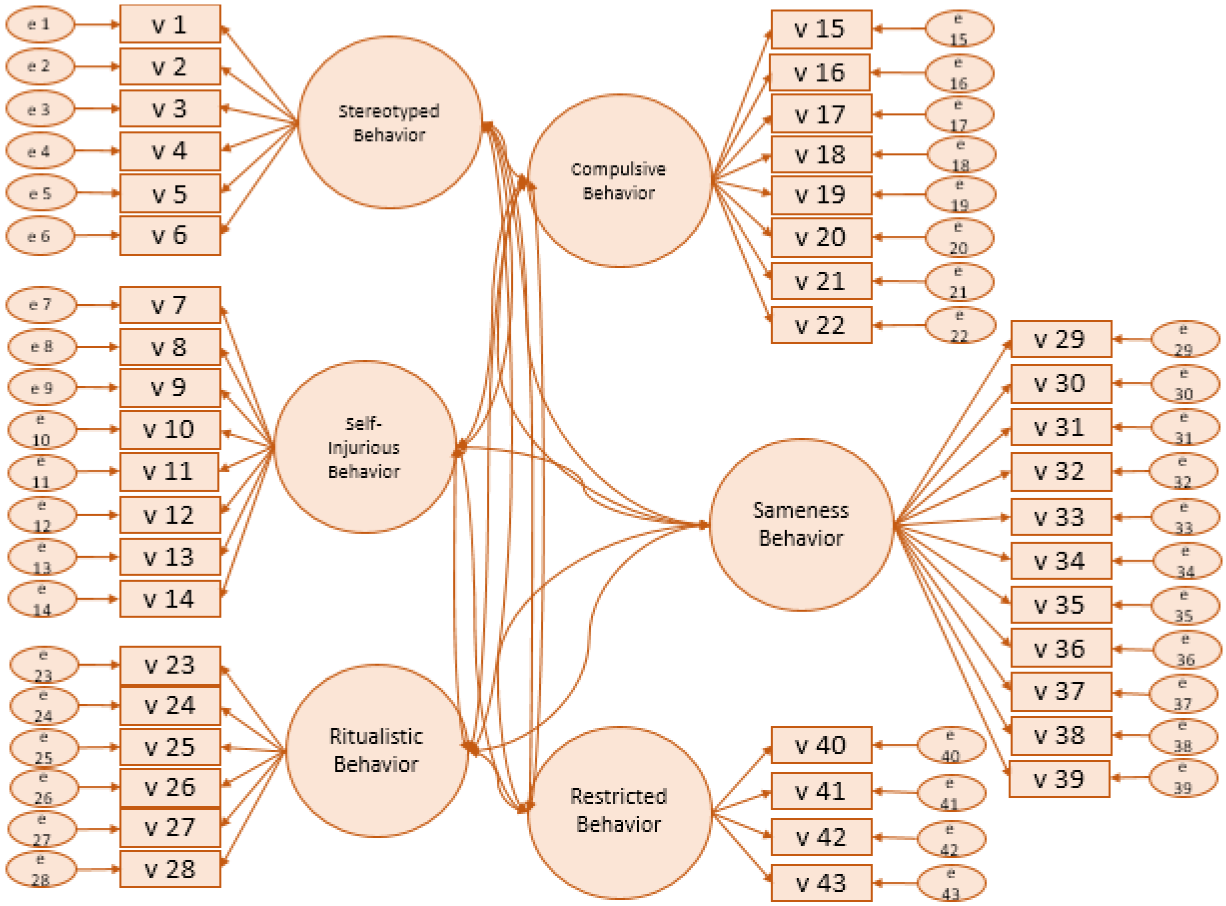


Figure C 2. Lam and Aman (2007) Five-Factor Model Path Diagram

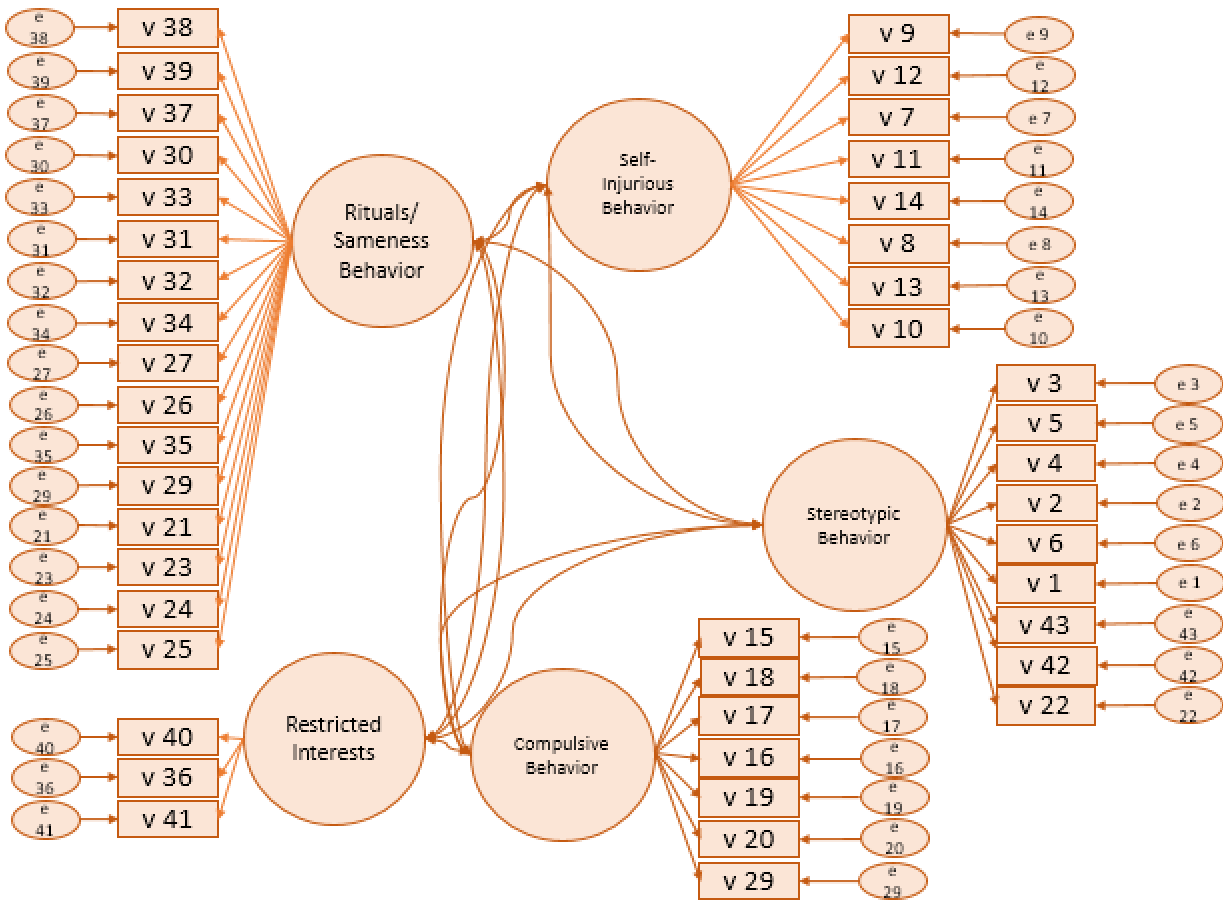


Figure C 3. Mirenda et al. (2010) Five-Factor Model Path Diagram

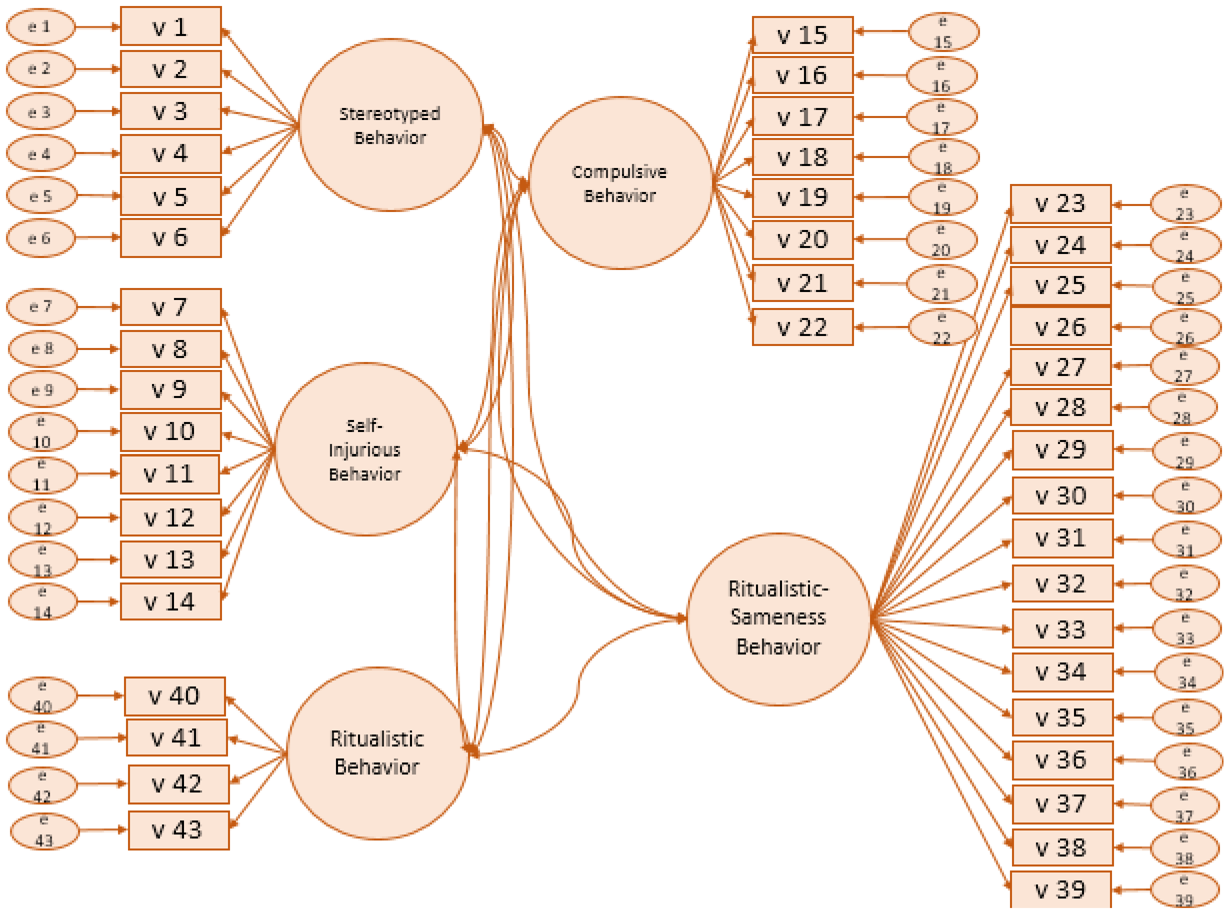


Figure C 4. Bishop et al. (2013) Five-Factor Model Path Diagram

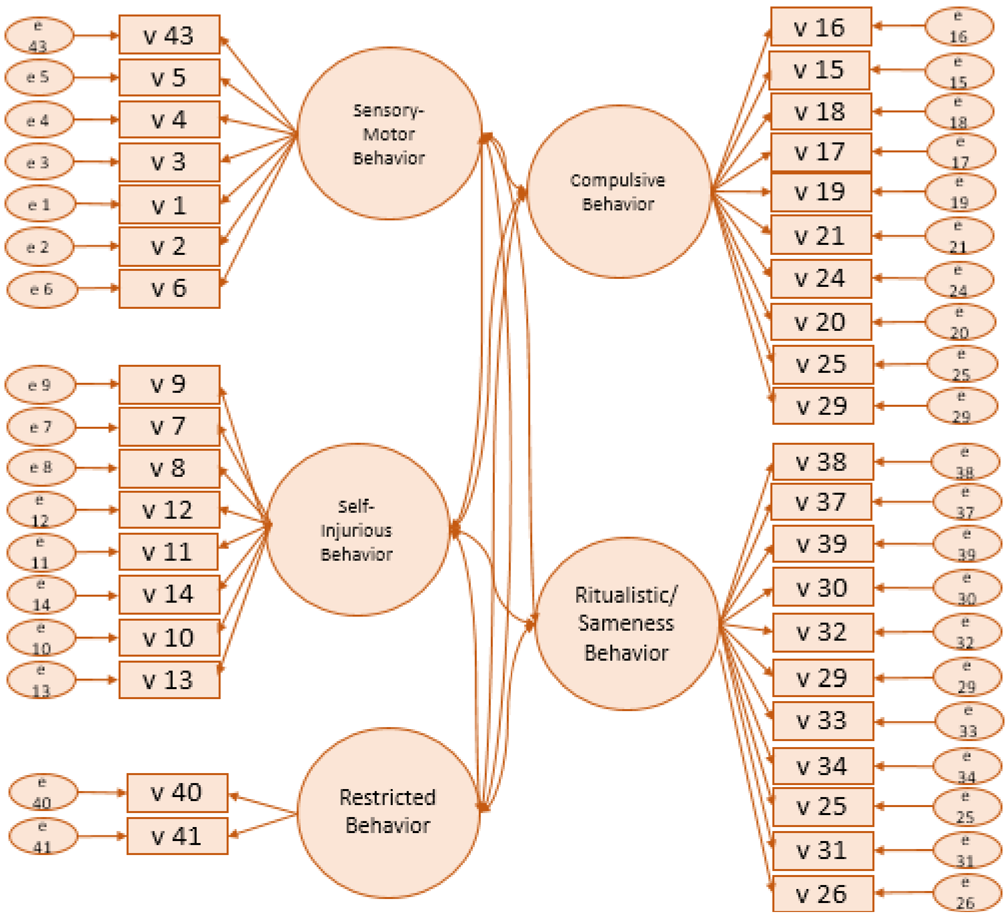


Figure C 5. Martinez-Gonzalez & Piqueras (2017) Six-Factor Model Path Diagram

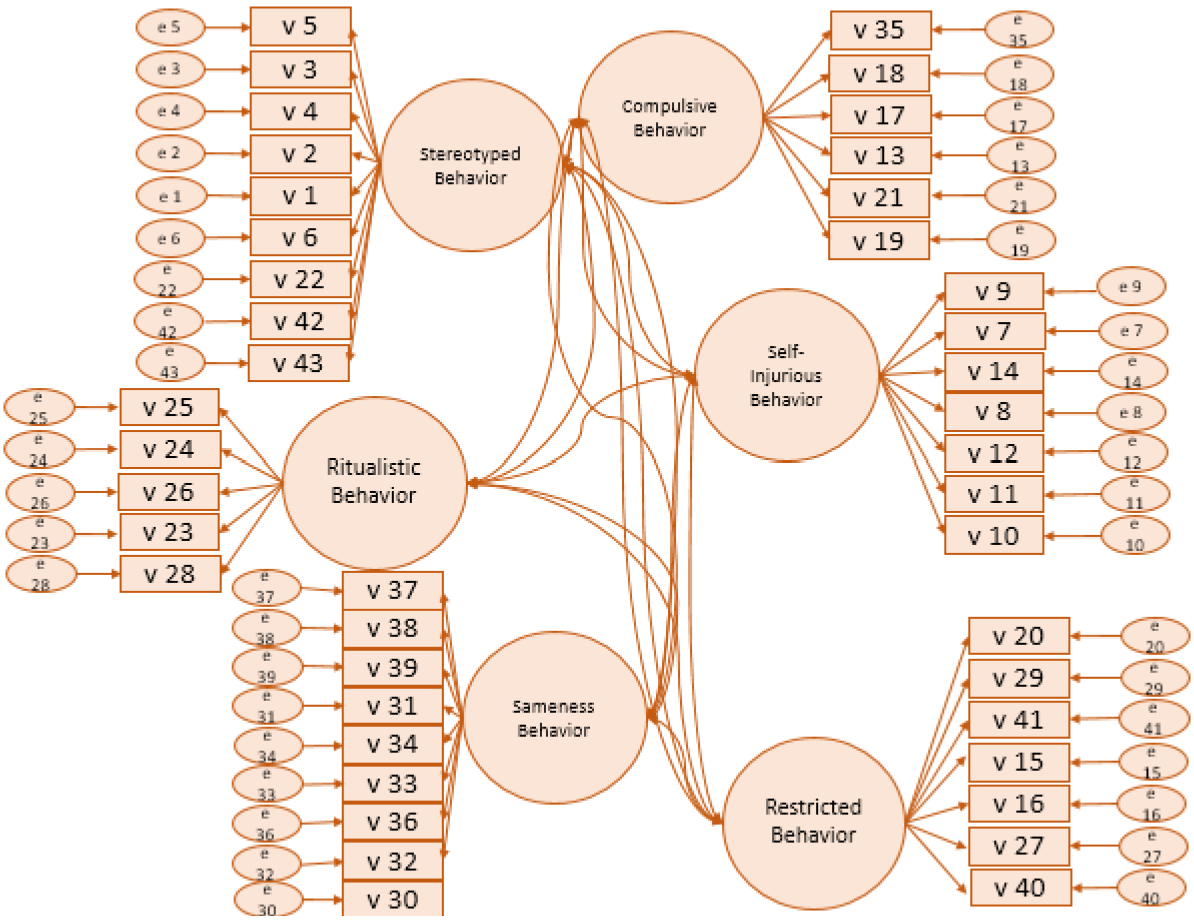


Figure C 6. Russell et al. (2019) Four-Factor Model Path Diagram

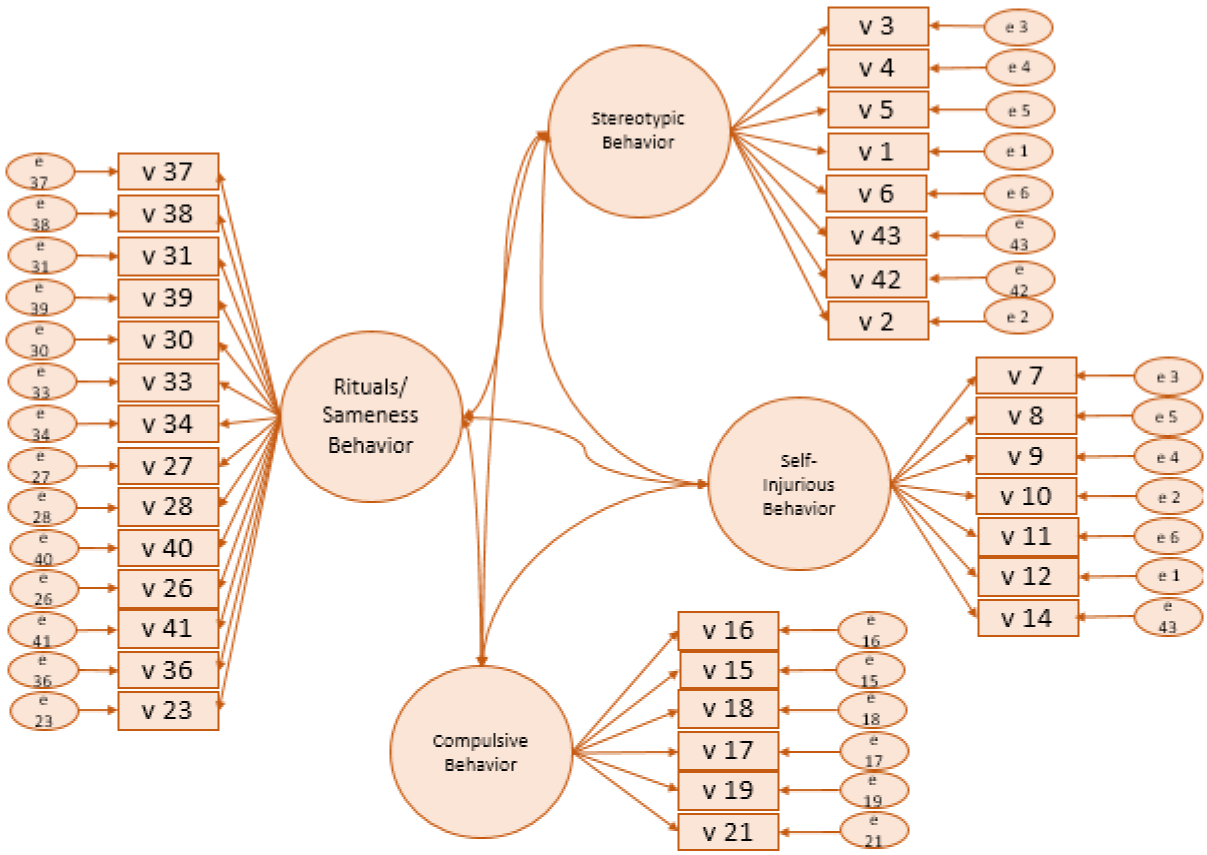
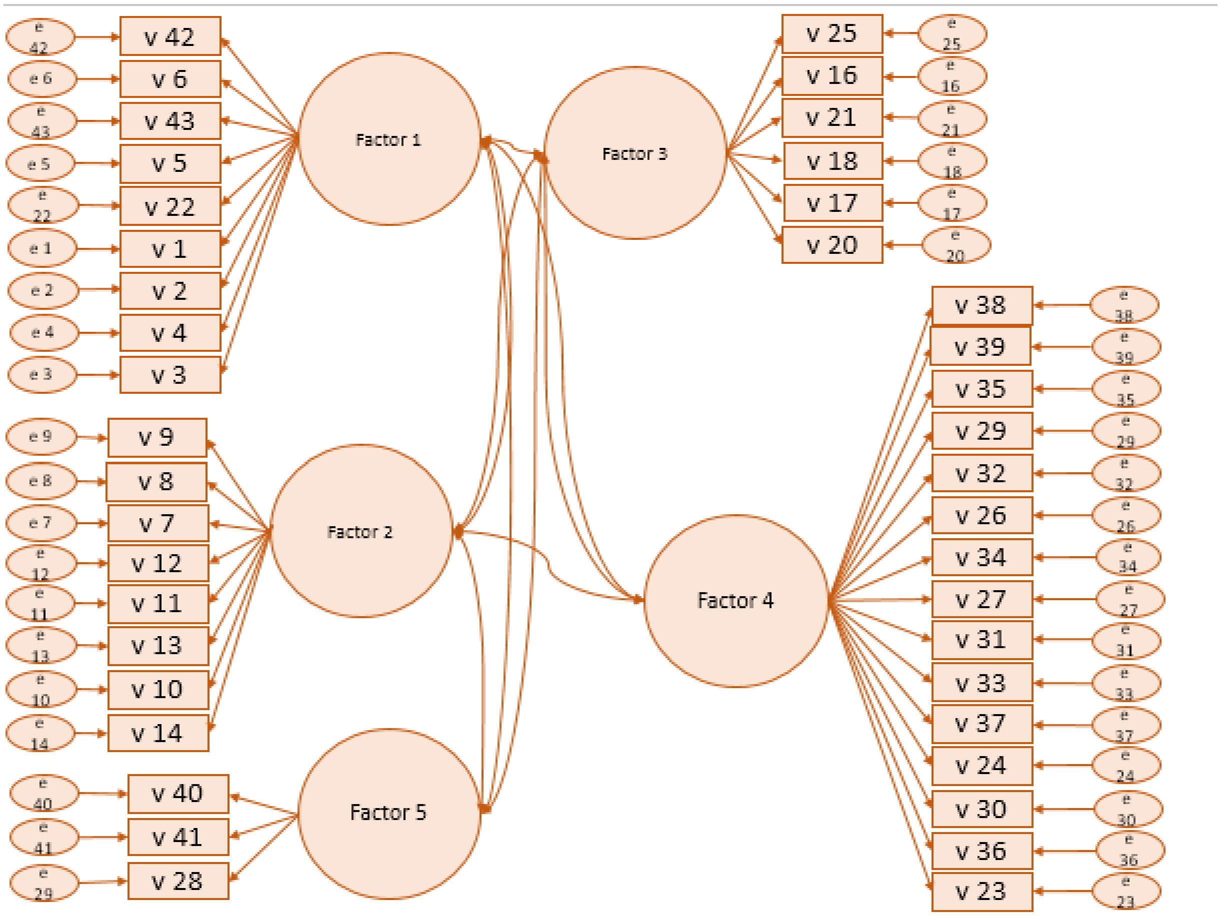


Figure C 7. Sturm et al. (2022) Five-Factor Model



APPENDIX D: Existing Model Study Two CFA Statistics

Table D 1. Bodfish et al. (2000) Six-Factor Model Parameter Estimates, Standard Errors, R^2, and Residual Variance

Factor	Item #	Item Stem	Parameter Estimate	Standard Error (S.E.)	Parameter Estimate/ Standard Error	R^2	Residual Variance
Stereotyped Behaviors	3	HAND/FINGER	0.758	0.041	18.459	0.575	0.425
	5	OBJECT USAGE	0.825	0.051	16.307	0.681	0.319
	4	LOCOMOTION	0.873	0.040	21.798	0.762	0.238
	1	WHOLE BODY	0.612	0.059	10.362	0.374	0.626
	2	HEAD	0.627	0.062	10.197	0.394	0.606
	6	SENSORY	0.877	0.038	12.948	0.769	0.626
Self-Injurious Behaviors	7	HITS SELF WITH BODY PART	0.798	0.045	17.852	0.638	0.362
	9	HITS SELF WITH OBJECT	0.879	0.047	18.878	0.773	0.227
	8	HITS SELF AGAINST SURFACE OR OBJECT	0.886	0.042	20.924	0.785	0.215
	10	BITES SELF	0.539	0.081	6.662	0.290	0.710
	11	PULLS	0.807	0.068	11.825	0.651	0.349
	12	RUBS OR SCRATCHES SELF	0.840	0.060	13.933	0.705	0.295
	13	INSERTS FINGER OR OBJECT	0.512	0.119	4.332	0.263	0.737
	14	SKIN PICKING	0.650	0.078	9.392	0.423	0.577
Compulsive Behaviors	21	REPEATING	0.828	0.049	16.897	0.686	0.314
	18	CHECKING	0.827	0.066	12.562	0.684	0.316
	15	ARRANGING/ ORDERING	0.860	0.029	29.404	0.739	0.261
	16	COMPLETENESS	0.873	0.032	26.875	0.762	0.238
	22	TOUCH/TAP	0.621	0.063	9.974	0.386	0.614
	17	WASHING/ CLEANING	0.654	0.079	8.311	0.428	0.572
	20	HOARDING/ SAVING	0.731	0.056	13.119	0.534	0.466

Table D1 (cont'd)

	19	COUNTING	0.858	0.064	13.472	0.737	0.466
Ritualistic Behavior	27	PLAY/LEISURE	0.861	0.031	27.946	0.741	0.259
	25	SELF-CARE - BATHROOM AND DRESSING	0.931	0.025	36.772	0.866	0.134
	24	SLEEPING/ BEDTIME	0.823	0.038	21.420	0.677	0.323
	23	EATING/ MEALTIME	0.782	0.039	20.121	0.611	0.389
	28	COMMUN-ICATION/SOCIAL INTERACTIONS	0.672	0.051	13.199	0.452	0.548
	26	TRAVEL/ TRANSPOR-TATION	0.936	0.026	36.725	0.877	0.123
Sameness Behaviors	39	Insists that specific things take place at specific times	0.913	0.024	38.208	0.834	0.166
	38	Insists on same routine, household, school, or work schedule everyday	0.882	0.025	35.430	0.778	0.222
	37	Resists changing activities; Difficulty with transitions	0.678	0.050	13.621	0.459	0.541
	35	Insists on using a particular door	0.933	0.055	17.027	0.871	0.129
	30	Objects to visiting new places	0.637	0.058	11.040	0.405	0.595
	33	Insists on sitting at the same place	0.694	0.056	12.477	0.482	0.518
	31	Becomes upset if interrupted in what he/she is doing	0.640	0.050	12.679	0.410	0.590
	34	Dislikes changes in appearance or behavior of the people around him/her	0.674	0.069	9.776	0.454	0.546

Table D1 (cont'd)

	36	Likes the same CD, tape, record, or piece of music played continually; Likes same move/video or part of movie /video	0.696	0.055	17.027	0.485	0.515
	29	Insists that things remain in the same place(s)	0.839	0.024	38.208	0.704	0.296
	32	Insists on walking in a particular pattern	0.807	0.050	16.017	0.651	0.349
Ritualistic Behaviors	40	Fascination, preoccupation with one subject or activity	0.826	0.036	22.950	0.682	0.318
	41	Strongly attached to one specific object	0.791	0.043	18.245	0.626	0.374
	42	Preoccupation with part(s) of object rather than the whole object	0.832	0.046	18.024	0.691	0.309
	43	Fascination, preoccupation with movement / things that move	0.821	0.046	17.797	0.675	0.325

Table D 2. Mirenda et al. (2010) Five-Factor Model Parameter Estimates, Standard Errors, R^2, and Residual Variance

Factor	Item #	Item Stem	Parameter Estimate	Standard Error (S.E.)	Parameter Estimate/ Standard Error	R^2	Residual Variance
Rituals-Sameness Behaviors	26	TRAVEL/ TRANSPOR-TATION	0.847	0.037	22.752	0.718	0.282
	27	PLAY/LEISURE	0.803	0.035	22.823	0.645	0.355
	28	COMMUN-ICATION/ SOCIAL INTERACTIONS	0.634	0.051	12.503	0.402	0.598
	30	Objects to visiting new places	0.639	0.057	11.176	0.408	0.592
	31	Becomes upset if interrupted in what he/she is doing	0.660	0.049	13.582	0.436	0.564
	32	Insists on walking in a particular pattern	0.791	0.052	15.325	0.626	0.374
	33	Insists on sitting at the same place	0.702	0.054	13.094	0.493	0.507
	34	Dislikes changes in appearance or behavior of the people around him/her	0.668	0.070	9.584	0.446	0.554
	35	Insists on using a particular door	0.896	0.052	17.093	0.803	0.197
	37	Resists changing activities; Difficulty with transitions	0.698	0.048	14.539	0.487	0.513
	38	Insists on same routine, household, school, or work schedule everyday	0.885	0.025	35.991	0.784	0.216
	39	Insists that specific things take place at specific times	0.927	0.023	40.755	0.859	0.141

Table D2 (cont'd)

Self-Injurious Behaviors	7	HITS SELF WITH BODY PART	0.801	0.042	19.094	0.641	0.359
	8	HITS SELF AGAINST SURFACE OR OBJECT	0.889	0.040	22.368	0.790	0.210
	9	HITS SELF WITH OBJECT	0.886	0.043	20.401	0.786	0.214
	10	BITES SELF	0.555	0.077	7.177	0.308	0.692
	11	PULLS	0.811	0.064	12.612	0.657	0.343
	12	RUBS OR SCRATCHES SELF	0.815	0.059	13.918	0.664	0.336
	13	INSERTS FINGER OR OBJECT	0.504	0.115	4.389	0.254	0.746
	14	SKIN PICKING	0.644	0.075	8.535	0.414	0.586
STB	1	WHOLE BODY	0.588	0.057	10.306	0.346	0.654
	2	HEAD	0.585	0.060	9.705	0.340	0.660
	3	HAND/FINGER	0.722	0.041	17.782	0.522	0.478
	4	LOCOMOTION	0.807	0.035	22.880	0.652	0.348
	5	OBJECT USAGE	0.778	0.045	17.128	0.605	0.395
	6	SENSORY	0.828	0.036	22.729	0.686	0.314
	22	TOUCH/TAP	0.700	0.063	11.122	0.490	0.510
	42	Preoccupation with part(s) of object rather than the whole object	0.790	0.044	17.837	0.623	0.377
	43	Fascination, preoccupation with movement / things that move	0.783	0.044	17.903	0.612	0.388
Compulsive Behaviors	15	ARRANGING/ ORDERING	0.857	0.031	27.838	0.734	0.266
	16	COMPLETENESS	0.899	0.034	26.569	0.809	0.191
	17	WASHING/ CLEANING	0.678	0.082	8.294	0.460	0.540

Table D2 (cont'd)

	18	CHECKING	0.855	0.066	12.965	0.731	0.269
	19	COUNTING	0.895	0.062	14.390	0.800	0.200
	20	HOARDING/ SAVING	0.755	0.055	13.796	0.570	0.430
Restricted Interests	36	Likes the same CD, tape, record, or piece of music played continually; Likes same move/video or part of movie /video	0.820	0.057	14.260	0.672	0.328
	40	Fascination, preoccupation with one subject or activity	0.808	0.035	23.192	0.653	0.347
	41	Strongly attached to one specific object	0.798	0.039	20.290	0.636	0.364

Table D 3. Martinez-Gonzalez & Piqueras (2017) Six-Factor Model Parameter Estimates, Standard Errors, R^2, and Residual Variance

Factor	Item #	Item Stem	Parameter Estimate	Standard Error (S.E.)	Parameter Estimate/ Standard Error	R^2	Residual Variance
Stereotyped Behaviors	5	OBJECT USAGE	0.778	0.047	16.681	0.605	0.395
	3	HAND/FINGER	0.716	0.043	16.837	0.513	0.487
	4	LOCOMOTION	0.815	0.036	22.632	0.665	0.335
	2	HEAD	0.579	0.063	9.213	0.336	0.664
	1	WHOLE BODY	0.566	0.059	9.535	0.320	0.680
	6	SENSORY	0.829	0.038	21.701	0.687	0.313
	22	TOUCH/TAP	0.712	0.063	11.307	0.508	0.492
	42	Preoccupation with part(s) of object rather than the whole object	0.792	0.046	17.085	0.628	0.372
	43	Fascination, preoccupation with movement / things that move	0.784	0.045	17.251	0.615	0.385
Self-Injurious Behaviors	9	HITS SELF WITH OBJECT	0.883	0.045	19.420	0.781	0.219
	7	HITS SELF WITH BODY	0.796	0.044	17.929	0.633	0.367
	14	SKIN PICKING	0.643	0.078	8.200	0.413	0.587
	8	HITS SELF AGAINST SURFACE OR OBJECT	0.887	0.040	21.990	0.413	0.213
	12	RUBS OR SCRATCHES SELF	0.835	0.059	14.115	0.787	0.303
	11	PULLS	0.803	0.068	11.890	0.645	0.355
	10	BITES SELF	0.537	0.081	6.589	0.288	0.712
Compulsive Behaviors	35	Insists on using a particular door	0.904	0.058	15.518	0.817	0.183
	18	CHECKING	0.798	0.069	11.585	0.637	0.363

Table D3 (cont'd)

Ritualistic Behavior	17	WASHING/ CLEANING	0.605	0.076	8.007	0.366	0.634
	13	INSERTS FINGER OR OBJECT	0.334	0.110	3.026	0.112	0.888
	21	REPEATING	0.800	0.048	16.694	0.640	0.360
	25	SELF-CARE – BATHROOM AND DRESSING	0.944	0.025	38.299	0.891	0.109
	24	SLEEPING/BEDTIME	0.834	0.037	22.432	0.696	0.304
	26	TRAVEL/ TRANSPORTATION	0.947	0.040	20.125	0.897	0.103
	23	EATING/ MEALTIME	0.800	0.025	37.268	0.640	0.360
Sameness Behaviors	37	Resists changing activities; Difficulty with transitions	0.689	0.050	13.820	0.474	0.526
	38	Insists on same routine, household, school, or work schedule everyday	0.895	0.025	36.143	0.802	0.198
	39	Insists that specific things take place at specific times	0.933	0.023	40.804	0.870	0.130
	31	Becomes upset if interrupted in what he/she is doing	0.672	0.048	13.935	0.452	0.548
	34	Dislikes changes in appearance or behavior of the people around him/her	0.683	0.071	9.573	0.466	0.534
	33	Insists on sitting at the same place	0.709	0.056	12.642	0.503	0.497
	36	Likes the same CD, tape, record, or piece of music played continually	0.725	0.055	13.160	0.526	0.474

Table D3 (cont'd)

	32	Insists on walking in a particular pattern	0.821	0.053	15.366	0.673	0.327
Ritualistic Behaviors	20	HOARDING/ SAVING	0.726	0.058	12.510	0.527	0.473
	29	Insists that things remain in the same place(s)	0.844	0.043	19.861	0.713	0.287
	41	Strongly attached to one specific object	0.686	0.050	13.642	0.470	0.530
	15	ARRANGING/ ORDERING	0.845	0.031	27.670	0.715	0.285
	16	COMPLETENESS	0.863	0.035	24.693	0.744	0.256
	27	PLAY/LEISURE	0.835	0.037	22.732	0.698	0.302
	40	Fascination, preoccupation with one subject or activity	0.705	0.042	16.910	0.497	0.503

Table D 4. Hooker et al. (2019) Five-Factor Model Parameter Estimates, Standard Errors, R^2, and Residual Variance

Factor	Item #	Item Stem	Parameter Estimate	Standard Error (S.E.)	Parameter Estimate/ Standard Error	R^2	Residual Variance
Stereotypic Behavior	1	WHOLE BODY	0.612	0.059	10.368	0.375	0.625
	2	HEAD	0.628	0.062	10.199	0.394	0.606
	3	HAND/FINGER	0.758	0.041	18.462	0.575	0.425
	4	LOCOMOTION	0.873	0.040	21.788	0.762	0.238
	5	OBJECT USAGE	0.825	0.051	16.312	0.681	0.319
	6	SENSORY	0.877	0.038	22.934	0.769	0.231
Self-Injurious Behavior	7	HITS SELF WITH BODY	0.798	0.045	17.855	0.637	0.363
	8	HITS SELF AGAINST SURFACE OR OBJECT	0.884	0.043	20.696	0.782	0.218
	9	HITS SELF WITH OBJECT	0.879	0.047	18.820	0.773	0.227
	10	BITES SELF	0.538	0.081	6.634	0.290	0.710
	11	PULLS	0.807	0.068	11.810	0.652	0.348
	12	RUBS OR SCRATCHES SELF	0.842	0.061	13.900	0.708	0.262
	13	INSERTS FINGER OR OBJECT	0.515	0.118	4.347	0.265	0.735
	14	SKIN PICKING	0.651	0.078	8.377	0.424	0.576
Compulsive Behavior	15	ARRANGING/ ORDERING	0.860	0.029	29.381	0.739	0.261
	16	COMPLETENESS	0.873	0.033	26.844	0.762	0.238
	17	WASHING/ CLEANING	0.654	0.079	8.316	0.427	0.573
	18	CHECKING	0.826	0.066	12.555	0.683	0.317
	19	COUNTING	0.858	0.064	13.439	0.737	0.263
	20	HOARDING/SAVING	0.731	0.056	13.116	0.534	0.573
	21	REPEATING	0.828	0.049	16.893	0.686	0.317
	22	TOUCH/TAP	0.620	0.063	9.788	0.385	0.263

Table D4 (cont'd)

Ritualistic-Sameness Behavior	23	EATING/MEALTIME	0.731	0.039	18.868	0.534	0.466
	24	SLEEPING/ BEDTIME	0.786	0.040	19.501	0.618	0.382
	25	SELF-CARE – BATHROOM AND DRESSING	0.899	0.027	33.646	0.808	0.192
	26	TRAVEL/ TRANSPORTATION	0.887	0.027	32.569	0.787	0.213
	27	PLAY/LEISURE	0.805	0.032	24.994	0.649	0.351
	28	COMMUNICATION/ SOCIAL INTERACTIONS	0.628	0.050	12.521	0.394	0.606
	29	Insists that things remain in the same place(s)	0.812	0.044	18.521	0.660	0.340
	30	Objects to visiting new places	0.619	0.058	10.742	0.383	0.617
	31	Becomes upset if interrupted in what he/she is doing	0.620	0.051	12.213	0.385	0.615
	32	Insists on walking in a particular pattern	0.782	0.051	15.223	0.611	0.389
	33	Insists on sitting at the same place	0.677	0.056	12.143	0.459	0.541
	34	Dislikes changes in appearance or behavior of the people around him/her	0.658	0.069	9.570	0.432	0.568
	35	Insists on using a particular door	0.885	0.053	16.700	0.783	0.217
	36	Likes the same CD, tape, record, or piece of music played continually; Likes same move/video or part of movie /video	0.675	0.055	12.343	0.456	0.544
	37	Resists changing activities; Difficulty with transitions	0.659	0.050	13.250	0.435	0.565

Table D4 (cont'd)

Restricted Behavior	38	Insists on same routine, household, school, or work schedule everyday	0.864	0.026	32.906	0.746	0.254
	39	Insists that specific things take place at specific times	0.900	0.025	36.203	0.810	0.190
	40	Fascination, preoccupation with one subject or activity	0.826	0.036	22.938	0.682	0.318
	41	Strongly attached to one specific object	0.791	0.043	18.229	0.626	0.374
	42	Preoccupation with part(s) of object rather than the whole object	0.832	0.046	18.020	0.691	0.309
	43	Fascination, preoccupation with movement / things that move	0.821	0.046	17.793	0.675	0.325

Table D 5. Sturm et al. (2019) Five-Factor Model Parameter Estimates, Standard Errors, R^2, and Residual Variance

Factor	Item #	Item Stem	Parameter Estimate	Standard Error (S.E.)	Parameter Estimate/ Standard Error	R^2	Residual Variance
Factor I: Stereotyped Behavior .	42	Preoccupation with part(s) of object rather than the whole object	0.798	0.046	17.473	0.638	0.362
	6	SENSORY	0.828	0.039	21.514	0.686	0.314
	43	Fascination, preoccupation with movement / things that move	0.791	0.045	17.452	0.625	0.375
	5	OBJECT USAGE	0.772	0.048	16.027	0.595	0.405
	22	TOUCH/TAP	0.719	0.064	11.306	0.517	0.483
	1	WHOLE BODY	0.561	0.061	9.243	0.315	0.685
	2	HEAD	0.573	0.064	8.972	0.328	0.672
	4	LOCOMOTION	0.816	0.036	22.610	0.666	0.334
	3	HAND/FINGER	0.712	0.043	16.564	0.507	0.493
Factor II: Self-Injurious Behavior	9	HITS SELF WITH OBJECT	0.880	0.046	18.947	0.774	0.226
	8	HITS SELF AGAINST SURFACE OR OBJECT	0.886	0.042	21.033	0.785	0.215
	7	HITS SELF WITH BODY	0.797	0.044	17.999	0.635	0.365
	12	RUBS OR SCRATCHES SELF	0.839	0.061	13.858	0.704	0.296
	11	PULLS	0.808	0.068	11.892	0.653	0.347
	13	INSERTS FINGER OR OBJECT	0.516	0.118	4.375	0.266	0.734
	10	BITES SELF	0.540	0.081	6.655	0.291	0.709
	14	SKIN PICKING	0.649	0.078	8.335	0.422	0.578
Factor III	25	SELF-CARE – BATHROOM AND DRESSING	0.946	0.029	32.630	0.895	0.105

Table D5 (cont'd)

	16	COMPLETENESS	0.849	0.034	25.056	0.721	0.279
	21	REPEATING	0.802	0.049	16.531	0.644	0.356
	18	CHECKING	0.801	0.068	11.852	0.642	0.358
	15	ARRANGING/ ORDERING	0.837	0.030	28.228	0.701	0.299
	19	COUNTING	0.831	0.063	13.100	0.691	0.309
	17	WASHING/ CLEANING	0.630	0.076	8.255	0.397	0.603
	20	HOARDING/ SAVING	0.709	0.056	12.696	0.503	0.497
Factor IV	38	Insists on same routine, household, school, or work schedule everyday	0.861	0.026	32.983	0.742	0.258
	39	Insists that specific things take place at specific times	0.897	0.025	36.032	0.805	0.195
	35	Insists on using a particular door	0.881	0.053	16.734	0.777	0.223
	29	Insists that things remain in the same place(s)	0.807	0.043	18.573	0.652	0.348
	32	Insists on walking in a particular pattern	0.779	0.052	15.092	0.606	0.394
	26	TRAVEL/ TRANSPOR-TATION	0.892	0.028	32.005	0.795	0.205
	34	Dislikes changes in appearance or behavior of the people around him/her	0.652	0.068	9.527	0.425	0.575
	27	PLAY/LEISURE	0.804	0.033	24.452	0.646	0.354
	31	Becomes upset if interrupted in what he/she is doing	0.619	0.051	12.228	0.384	0.616
		Insists on sitting at the same place	0.676	0.055	12.212	0.457	0.543

Table D5 (cont'd)

	37	Resists changing activities; Difficulty with transitions	0.658	0.050	13.258	0.433	0.567
	24	SLEEPING/ BEDTIME	0.788	0.041	19.039	0.622	0.378
	30	Objects to visiting new places	0.617	0.057	10.774	0.381	0.319
	36	Likes the same CD, tape, record, or piece of music played continually; Likes same move/video or part of movie /video	0.674	0.055	12.370	0.455	0.545
	23	EATING/ BEDTIME	0.729	0.039	18.642	0.531	0.469
Factor V	41	Strongly attached to one specific object	0.776	0.040	19.255	0.603	0.397
	40	Fascination, preoccupation with one subject or activity	0.813	0.034	24.022	0.661	0.339
	28	COMMUN-ICATION/ SOCIAL INTERACTION	0.750	0.060	12.520	0.562	0.438

APPENDIX E: Lam and Aman (2007) Five-Factor Model Ancillary Tables

Table E1. Lam and Aman (2007) Five-Factor Solution Published in Lam (2004)

	Factor I	**Factor II**	**Factor III**	**Factor IV**	**Factor V**
1. Body movements	0.125	0.116	**0.470**	0.036	-0.286
2. Head movements	0.105	0.073	**0.512**	0.023	-0.134
3. Finger movements	0.043	0.057	**0.651**	-0.004	-0.086
4. Locomotion	-0.037	0.113	**0.566**	0.048	0.072
5. Object usage	-0.023	0.182	**0.628**	-0.060	0.168
6. Sensory	0.053	0.178	**0.492**	0.022	0.100
8. Hits against surface	0.044	**0.551**	0.169	-0.034	0.055
9. Hits w/ object	0.014	**0.703**	0.070	-0.015	0.029
10.Bites self	0.088	**0.506**	0.129	-0.142	-0.026
11. Pulls hair/skin	0.123	**0.623**	-0.053	0.051	-0.052
12. Rubs/scratches	-0.002	**0.683**	-0.018	0.040	0.033
13. Inserts finger/object	-0.003	**0.511**	0.086	0.057	-0.025
14. Picks skin	0.030	**0.570**	-0.260	0.152	-0.050
16. Completeness	-0.114	0.174	0.111	**0.545**	0.125
17. Washing	-0.066	0.165	0.015	**0.553**	-0.026
18. Checking	0.084	0.096	0.077	**0.577**	-0.197
19. Counting	-0.049	0.105	0.056	**0.521**	0.135
20. Hoarding	0.080	0.135	-0.078	**0.466**	0.119
21. *Repeating*	*0.351*	*0.018*	*0.130*	*0.275*	*0.101*
22. Needs to touch/tap	0.147	0.040	**0.351**	0.110	0.116
24. *Sleeping/bedtime*	*0.319*	*-0.044*	*0.012*	*0.363*	*0.176*
25. *Self-care routine*	*0.427*	*-0.066*	*-0.002*	*0.410*	*0.036*
26. Transportation routine	**0.479**	-0.057	0.112	0.286	0.023
27. Play/leisure routine	**0.511**	-0.028	0.077	0.180	0.161
28. Communication	**0.373**	0.063	-0.234	0.096	0.244
30. No new places	**0.712**	0.002	0.000	-0.057	-0.081
31. No interruption	**0.573**	0.132	-0.013	0.078	0.130
32. Walks certain way	**0.566**	0.093	0.199	0.044	-0.184
33. Sits certain place	**0.631**	-0.032	0.201	0.114	-0.172
34. Appearance/behavior of others	**0.526**	0.114	0.104	0.088	0.010
35. Uses certain door	**0.477**	-0.017	0.235	0.070	-0.018
36. Videotapes	0.297	0.003	-0.001	0.025	**0.492**
37. Difficult transitions	**0.740**	0.060	-0.018	-0.133	0.184
38. Insists on routine	**0.780**	0.081	-0.014	-0.011	0.076
39. Insists on time	**0.770**	0.089	-0.087	0.031	0.076
40. Preoccupation with subject	0.138	0.017	0.078	0.082	**0.646**
41. Attached to object	0.127	0.058	0.313	0.135	**0.446**
42. Preoccupied with part of object	0.079	0.014	**0.444**	0.142	0.338
43. Preoccupation with movement	0.027	0.015	**0.458**	0.168	0.292

Table E 2. Re-constructed Lam and Aman (2007) Five-Factor Model Five-Factor Solution

	Factor I	**Factor II**	**Factor III**	**Factor IV**	**Factor V**
1. Body movements	0.125	0.116	**0.470**	0.036	-0.286
2. Head movements	0.105	0.073	**0.512**	0.023	-0.134
3. Finger movements	0.043	0.057	**0.651**	-0.004	-0.086
4. Locomotion	-0.037	0.113	**0.566**	0.048	0.072
5. Object usage	-0.023	0.182	**0.628**	-0.060	0.168
6. Sensory	0.053	0.178	**0.492**	0.022	0.100
8. Hits against surface	0.044	**0.551**	0.169	-0.034	0.055
9. Hits w/ object	0.014	**0.703**	0.070	-0.015	0.029
10.Bites self	0.088	**0.506**	0.129	-0.142	-0.026
11. Pulls hair/skin	0.123	**0.623**	-0.053	0.051	-0.052
12. Rubs/scratches	-0.002	**0.683**	-0.018	0.040	0.033
13. Inserts finger/object	-0.003	**0.511**	0.086	0.057	-0.025
14. Picks skin	0.030	**0.570**	-0.260	0.152	-0.050
16. Completeness	-0.114	0.174	0.111	**0.545**	0.125
17. Washing	-0.066	0.165	0.015	**0.553**	-0.026
18. Checking	0.084	0.096	0.077	**0.577**	-0.197
19. Counting	-0.049	0.105	0.056	**0.521**	0.135
20. Hoarding	0.080	0.135	-0.078	**0.466**	0.119
21. Repeating	**0.351**	0.018	0.130	0.275	0.101
22. Needs to touch/tap	0.147	0.040	**0.351**	0.110	0.116
24. Sleeping/bedtime	0.319	-0.044	0.012	**0.363**	0.176
25. Self-care routine	**0.427**	-0.066	-0.002	0.410	0.036
26. Transportation routine	**0.479**	-0.057	0.112	0.286	0.023
27. Play/leisure routine	**0.511**	-0.028	0.077	0.180	0.161
28. Communication	**0.373**	0.063	-0.234	0.096	0.244
30. No new places	**0.712**	0.002	0.000	-0.057	-0.081
31. No interruption	**0.573**	0.132	-0.013	0.078	0.130
32. Walks certain way	**0.566**	0.093	0.199	0.044	-0.184
33. Sits certain place	**0.631**	-0.032	0.201	0.114	-0.172
34. Appearance/behavior of others	**0.526**	0.114	0.104	0.088	0.010
35. Uses certain door	**0.477**	-0.017	0.235	0.070	-0.018
36. Videotapes	0.297	0.003	-0.001	0.025	**0.492**
37. Difficult transitions	**0.740**	0.060	-0.018	-0.133	0.184
38. Insists on routine	**0.780**	0.081	-0.014	-0.011	0.076
39. Insists on time	**0.770**	0.089	-0.087	0.031	0.076
40. Preoccupation with subject	0.138	0.017	0.078	0.082	**0.646**

Table E2 (cont'd)

41. Attached to object	0.127	0.058	0.313	0.135	**0.446**
42. Preoccupied with part of object	0.079	0.014	**0.444**	0.142	0.338
43. Preoccupation with movement	0.027	0.015	**0.458**	0.168	0.292