American Journal on Intellectual and Developmental Disabilities Vineland-3 Structural Validity and Interpretability of Domain Scores: Implications for Practitioners

--Manuscript Draft--

Manuscript Number:	AJIDD-D-19-00008R1
Article Type:	Research Report
Keywords:	Vineland-3; Intellectual Disability; Disability Assessment; Reliability; Validity
Corresponding Author:	Vincent Pandolfi, Ph.D. Rochester Institute of Technology Rochester, NY UNITED STATES
First Author:	Vincent Pandolfi, Ph.D.
Order of Authors:	Vincent Pandolfi, Ph.D.
	Caroline Ilona Magyar, Ph.D.
Manuscript Region of Origin:	UNITED STATES
Abstract:	The Vineland-3 (Sparrow et al., 2016) purports to measure three dimensions of adaptive behavior but empirical evidence pertaining to its structural validity is lacking. In this study factor analyses were conducted on the standardization sample data for the Comprehensive forms within the 11 to 20-year-old age range. Results did not support the three Domain structure of the test and indicated Domain scores did not add additional information about an individual's adaptive performance that was not already accounted for by the Adaptive Behavior Composite (ABC) score alone. Practitioners assessing adolescents with developmental conditions should consider using the ABC score within a multimethod assessment protocol for the various purposes of adaptive behavior assessment including the identification of Intellectual Disability.

Vineland-3 Structural Validity and Interpretability of Domain Scores: Implications for Practitioners Assessing Adolescents with Developmental Conditions

Abstract

The Vineland-3 (Sparrow et al., 2016) purports to measure three dimensions of adaptive behavior but empirical evidence pertaining to its structural validity is lacking. In this study factor analyses were conducted on the standardization sample data for the Comprehensive forms within the 11 to 20-year-old age range. Results did not support the three Domain structure of the test and indicated Domain scores did not add additional information about an individual's adaptive performance that was not already accounted for by the Adaptive Behavior Composite (ABC) score alone. Practitioners assessing adolescents with developmental conditions should consider using the ABC score within a multimethod assessment protocol for the various purposes of adaptive behavior assessment including the identification of Intellectual Disability.

Key Words: Vineland-3, Intellectual Disability, Disability Assessment, Reliability, Validity

Vineland-3 Structural Validity and Interpretability of Domain Scores: Implications for Practitioners Assessing Adolescents with Developmental Conditions

A number of tests are available for use in adaptive behavior assessment (see Price et al., 2018) and one commonly used is the Vineland Adaptive Behavior Scales-3rd Edition (Vineland-3; Sparrow et al., 2016). The Vineland-3 purports to measure multiple dimensions of adaptive behavior which the test authors define as "the performance of daily activities required for personal and social self-sufficiency" (Sparrow et al., 2016, p.10). The constructs it intends to measure appear consistent with the definition of adaptive behavior specified by the American Association on Intellectual and Developmental Disabilities (AAIDD, 2010) and the Diagnostic and Statistical Manual of Mental Disorders -5th Edition (DSM-5; American Psychiatric Association, 2013). The AAIDD defines adaptive behavior as set of conceptual, social, and practical skills that one learns and applies in everyday life. Sparrow et al. (2016) indicate that Vineland-3 data can be used when assessing individuals with developmental conditions for the purposes of diagnosing Intellectual Disability, intervention and support planning, and progress monitoring.

The AAIDD (2010) provides six technical standards that address the psychometric and administrative aspects of test selection when assessing an individual with a developmental condition. These standards emphasize the importance of using a reliable and valid test designed to measure adaptive behavior within a multimethod assessment protocol in order to capture multiple dimensional aspects of individual performance. While the Vineland-3 technical manual provides conceptual and empirical information important for establishing the test's reliability and validity, it does not contain evidence pertaining to the test's factor structure. The present study analyzed the Vineland-3 factor structure to determine the extent to which it reliably measures

multiple dimensions of adaptive behavior. Results will inform practitioners on how Vineland-3 data can be used as part of a multimethod approach to adaptive behavior assessment.

The Vineland-3 test structure is consistent with that of the Vineland-2 for which the test authors provided factor analytic support (see Sparrow et al., 2005). Like its predecessor, the Vineland-3 provides three levels of norm-referenced scores. The Adaptive Behavior Composite (ABC) reflects the test's total score and quantifies overall adaptive performance. Its Communication, Daily Living Skills, and Socialization Domain scores are purported to quantify performance within more narrow dimensions of adaptive behavior. There are also nine Subdomain scores that are purported to measure specific aspects of performance and three Subdomains are nested within each Domain. The Vineland-3's technical manual contains specific instructions on how these scores can be used for diagnosis, intervention planning, and progress monitoring. However, factor analytic evidence supporting the Vineland-3's three Domain structure is needed to help substantiate the appropriate use of these scores.

The current study utilized exploratory factor analyses (EFA) for two major purposes. The first purpose was to identify the number and nature of the constructs measured by the Vineland-3. The second purpose was to evaluate the interpretability of Domain scores which was done through the use of exploratory bifactor analysis (E-BFA). Although it is not a new method, E-BFA is becoming a more common complement to traditional exploratory factor analysis in contemporary psychometric research. Such is illustrated in recent evaluations of intelligence tests that, like adaptive behavior tests, are frequently used to assess individuals with developmental conditions (e.g., see Canivez et al., 2016).

[Insert Figure1 here]

A bifactor model features two kinds of factors, a general factor and group factors, all of which are orthogonal to one another. Figure 1 shows a bifactor model of the Vineland. The General factor (ABC) directly influences all nine Subdomains. There are three group factors (Domains) that directly influence their respective Subdomains. The Vineland-3 General factor (ABC) reflects what all the Subdomains have in common and it accounts for a portion of the variance in the Subdomain scores. Each group factor (Domain) reflects only what its Subdomains have in common. A group factor (Domain) accounts for additional variance within its Subdomains that was not accounted for by the General factor (ABC). E-BFA is well-suited for one of the major purposes of the present study (see Reise et al., 2010). It can be used to evaluate how much variance in Subdomain scores is accounted for by the General factor (ABC) and how much is accounted for by the group factors (Domains). Bifactor results can also be used to calculate model-based reliability coefficients which can inform users of the Vineland-3 about the extent to which Domain scores provide reliable information about an individual's adaptive performance that was not already accounted for by the ABC score alone.

The analyses used in this study were applied to the Comprehensive Interview and Parent/Caregiver Report forms for the 12 to 20-year-old age range, and the Comprehensive Teacher Report form for the 11 to 18-year-old age range. We focused on these age ranges because adolescence is a developmental period characterized by physical, social-emotional, behavioral, and contextual (e.g., expectations, environmental) changes (e.g., Borthwick-Duffy, 2007; XXXX et al., 2006) and reassessment of adaptive performance is often necessary for diagnostic decision-making and treatment planning. Data from these analyses can inform practitioners how the Vineland-3 fits into a multimethod assessment protocol when assessing adolescents with developmental conditions.

Method

Instrument

The Vineland-3 is a norm-referenced test of adaptive behavior developed for individuals birth to 90+ years of age. The technical manual describes the standardization and norming procedure and indicates that within individual age bands the standardization sample's demographic characteristics are consistent with 2014 United States census data across education level, race/ethnicity, and geographic region. For individuals aged 3 to 21 years the standardization sample is reportedly representative of people with disabilities across disability categories specified by the Individuals with Disabilities Education Improvement Act (2004). The percentage of people falling within each disability category was consistent with data from the National Center for Education Statistics (2014).

Vineland-3 test items purportedly assess three dimensions of adaptive behavior. These dimensions are labeled the Communication, Daily Living Skills, and Socialization Domains. Each Domain consists of three Subdomains as follows: (a) the Communication Domain includes the Receptive, Expressive, and Written Subdomains; (b) the Daily Living Skills Domain includes Subdomains labeled Personal, Domestic (Interview and Parent/Caregiver forms), Numeric (Teacher form), Community (Interview and Parent/Caregiver forms), and School/Community (Teacher form); and (c) the Socialization Domain includes the Interpersonal Relationships, Play and Leisure Skills, and Coping Skills Subdomains.

The Vineland-3 provides norm-referenced scores for the nine Subdomains, the three Domains, and the Adaptive Behavior Composite (ABC). The Subdomain v-scale scores have a mean of 15 and standard deviation of 3 (range 1 to 24). The Domains and ABC have a mean of 100 and standard deviation of 15 (range 20 to 140). Each Subdomain v-scale score is derived through the

6

raw score sum of its items' scores. Each Domain standard score is derived through the sum of its three Subdomains' v-scale scores. The ABC standard score is derived from the sum of the three Domain standard scores.

The Vineland-3 technical manual contains reliability data for the standardization sample. This includes internal consistency estimates (Coefficient alpha). For each of the Vineland-3 forms, the technical manual presents data for separate narrow age bands as well as the mean internal consistency across all age ranges (see Tables 6.1, 6.7, and 6.12 in the technical manual). We present a brief summary for those age ranges relevant to this study. Median internal consistency coefficients for the ABC ranged from .980 (Teacher form) to .990 (Parent/Caregiver form). For the Domains, median internal consistency ranged from .940 (Daily Living Skills, Teacher) to .980 (Communication, Parent/Caregiver; Socialization, Parent/Caregiver and Teacher). Median internal consistency for the Subdomains on the Interview form ranged from .870 (Receptive and Personal) to .965 (Coping). For the Parent/Caregiver form the median Coefficient alpha ranged from .940 (Personal) to .970 (Coping). Test-retest and internater reliability data are also presented in the manual, which the test authors cite as good to excellent.

Validity evidence based on test content, structure, special subgroups, and relationship of Vineland-3 scores to other measures was reported. Evidence for test content and structure is most relevant to the present study. Vineland-3 item content was based on the AAIDD and DSM-5 definitions of adaptive behavior. Its structure, which features the division of test items into three major Domains, is consistent with a multidimensional conceptualization of adaptive behavior. The subdivision of each Domain into three Subdomains seems to allow for a more specific assessment of narrow areas of adaptive performance. However, no exploratory or confirmatory factor analyses were reported to empirically evaluate the three Domain structure of the Vineland-3.

Data Analyzed

Exploratory factor analyses and exploratory bifactor analyses were conducted on the intercorrelations among the nine Subdomains for the Comprehensive Interview (ages 12-20, N = 480), Parent/Caregiver (ages 12-20, N = 480), and Teacher Report (ages 11 to 18, N = 500) forms. These correlation matrices appear in Tables 5.7, 5.8, and 5.9 of the Vineland-3 manual, respectively.

Data Analysis

Exploratory factor analysis (EFA). We used EFA to understand the number and nature of the constructs measured by each Vineland-3 form. All EFAs were performed with the Comprehensive Exploratory Factor Analysis program (Browne et al., 2008) using Maximum Likelihood factor extraction with oblique Geomin rotation. For each of the three forms we requested one, two, and three factors to be extracted. Limiting the number of extracted factors to three was indicated by the recommendation that a meaningful factor should be defined by at least two variables (Tabachnik & Fidell, 2007) although at least three would be preferred consistent with the current Vineland-3 test structure.

Acceptability of a factor model was informed by the conceptual meaningfulness of factors and several pieces of empirical evidence. First, each factor model was inspected for out-of-range parameter estimates which indicate an improperly specified model. Second, we examined the Root Mean Square Error of Approximation model fit statistic (RMSEA; Steiger & Lind, 1980). Values <.05 indicate close fit, values between .05 and .08 indicate acceptable fit, values between .08 and .10 indicate marginal fit, and values >.10 indicate poor fit (Browne & Cudeck, 1989).

We also examined the Expected Cross Validation Index (ECVI) which estimates how well a factor model might generalize to other samples. Although no absolute interpretive guidelines exist, smaller values are preferred (Fabrigar et al., 1999). Use of the ECVI involves comparing obtained values across competing factor models, and the model exhibiting the lowest ECVI is the preferred model. We considered an individual factor interpretable if it contained at least three Subdomains with a salient factor pattern loading. Although researchers vary in the use of a specific cutoff, we deemed a pattern coefficient $\geq .32$ as a reasonable lower limit (see Osborne, 2014; Tabachnick & Fidell, 2007). For conceptual clarity and parsimony, we also required that all Subdomains had a salient loading on only one factor. For multi-factor models the magnitude of factor correlations was examined: moderate to high correlations (e.g., \geq .40) suggested the presence of an unmodeled General factor that all Subdomains had in common and empirically justified the use of bifactor analysis (see Reise et al., 2007). Finally, the percentage of total variance accounted for by each factor model was examined. For the oblique multifactor models, it was not possible to unambiguously determine the percentage of variance accounted for by each factor alone (Pedhazur & Schmelkin, 1991; Tabachnik & Fidell, 2007).

Exploratory bifactor analysis (E-BFA). For each of the Vineland-3 forms two different models were analyzed using E-BFA methods: (a) the most acceptable model to emerge from the EFAs, and (b) the Vineland-3 model as it is currently structured. For E-BFA we utilized Maximum Likelihood factor extraction and orthogonal target bifactor rotation (e.g., see Reise, 2012). Target bifactor rotation was performed so that each Subdomain: (a) had a freely estimated factor loading on the General factor, (b) a freely estimated factor loading on only one group factor, and (c) zero or near-zero factor loadings on the remaining factors. For models that emerged from EFA, a Subdomain was assigned to the group factor on which it exhibited a salient

pattern loading in the EFA (\geq .32). If a Subdomain exhibited a salient loading on more than one factor in an EFA, it was assigned to the group factor on which it had its highest loading. For the E-BFAs we referenced Comrey and Lee's (1992) recommendation that orthogonal factor loadings > .55 be considered good indicators of a factor. The mutually orthogonal factors allowed for examination of the total and common variance in Vineland-3 data accounted for by the General factor and by each group factor. The explained common variance (ECV) statistic indicated the amount of common variance that was accounted for by a factor. As a ratio of the variance accounted for by a factor divided by the variance accounted for by all factors, it is considered an index of factor strength. As applied to the General factor, when ECV > .70 a test can be considered essentially unidimensional (see Bonifay et al., 2015; Rodriguez et al., 2016). To the extent ECV for the General factor was \geq .70 this indicated that the Vineland-3 Comprehensive Interview, Parent/Caregiver, and/or Teacher forms was primarily measuring one general adaptive behavior construct. As potential measurement scales, the computation of the total variance and ECV accounted for by each of the group factors was based only on those Subdomains which had salient loadings on the group factor in the EFAs. These variances were computed using the Omega statistical program (Watkins, 2013).

Omega coefficients. E-BFA results were used to compute model-based reliability statistics using the Omega program. The omega indices have been reviewed in detail elsewhere (e.g., see Reise et al., 2010; Rodriguez et al., 2016) and are summarized here. Coefficient omega (GO) reflects the proportion of variance in a scale attributable to all sources of common variance and is similar to Coefficient alpha. Omega hierarchical (GO_H) indicates the proportion of reliable variance uniquely accounted for by the General factor (e.g., ABC). Omega hierarchical subscale (GO_{HS}) reflects the proportion of reliable variance uniquely accounted for by a group factor (e.g., Domain); that is, after controlling for the effect of the General factor. To the extent G_{H} is high (e.g., $\geq .80$) and G_{HS} is relatively low, then most of the reliable variance in a test is accounted for by a General factor (e.g., see Rodriguez et al., 2016). For EFA-based models G_{HS} was based on those Subdomains which had salient loadings on the group factor in the EFAs.

Results

Exploratory Factor Analysis

Interview form. Table 1 contains EFA results for the Comprehensive Interview form.

[Insert Table 1 here]

The correlated three-factor model accounted for 60.4% of the total variance in Vineland-3 scores and evidenced good statistical fit with RMSEA = .060. It also had the lowest ECVI (.206) as compared to the one- (.518) and two-factor (.333) models. The three Communication and three Socialization Subdomains all had salient pattern coefficients on Factor I. Factor II featured salient pattern coefficients by each of the three Daily Living Skills Subdomains. Only two Subdomains had marginally salient pattern coefficients on Factor III, Written and Interpersonal Relationships. With only two relatively small coefficients, one of which was negative (Written), Factor III was a weak factor and difficult to interpret. Finally, the rather large correlation between Factors I and II (.801) suggested the presence of an unmodeled General factor that all nine Subdomains have in common. Thus, the conceptual meaningfulness of the three-factor model was suspect.

The correlated two-factor model accounted for 56.0% of the variance and demonstrated marginal statistical fit with RMSEA = .099. Salient coefficients on Factor I included Receptive, Expressive, and all three Socialization Subdomains. The Written Subdomain and all three Daily Living Skills Subdomains had salient coefficients on Factor II. Factors I and II correlated to a

substantial degree (.722) which once again indicated that all nine Subdomains shared an unmodeled General factor.

Finally, the one-factor model with an RMSEA = .120 demonstrated poor statistical fit. This suggested that additional factors were needed to better account for the correlations among the nine Subdomains. However, the single factor did account for 50.7% of the total variance and all factor loadings were very good to excellent indicators of the factor.

Parent/Caregiver form. Table 2 contains the EFA results for the Comprehensive Parent/Caregiver form.

[Insert Table 2 here]

The correlated three-factor model accounted for 65.9% of the variance and evidenced good statistical fit with RMSEA = .066. It had the lowest ECVI (.209) relative to the two- (.427) and one-factor (.671) models. Factor I contained the Receptive and Written Subdomains along with the three Daily Living Skills Subdomains. Factor II contained the three Socialization Subdomains. Factor III was a weak factor which contained only two Subdomains with salient coefficients, Receptive and Expressive. Because the Receptive Subdomain had coefficients of about equal magnitude on Factors I and III, it is difficult to understand what aspects of the construct represented by the Receptive Subdomain was related to it joining the Expressive Subdomain to form a separate factor. Factor correlations ranged from .391 (I and III) to .701 (I and II) which again suggested the presence of a General factor.

The correlated two-factor model (RMSEA = .121) and the one-factor model (RMSEA = .141) demonstrated poor statistical fit. The two-factor model accounted for 59.3% of the variance and evidenced a rather high factor correlation (.724) which once again suggested the presence of an unmodeled General factor. The one-factor model accounted for 53.9% of the variance and the

magnitude of factor loadings indicated that all nine Subdomains were very good to excellent indicators of the factor.

Teacher form. Table 3 contains EFA results for the Comprehensive Teacher form.

[Insert Table 3 here]

The correlated three-factor model accounted for 69.4% of the variance and evidenced acceptable statistical fit with RMSEA = .068. It had the lowest ECVI (.219) relative to the two-factor (.274) and one-factor (.788) models. This factor model reflected a Heywood case: 100% of the variance in School/Community was accounted for by the three factors which suggested an improper model. Factor I contained the Communication and Daily Living Skills Subdomains, as well as Interpersonal Relationships. Factor II contained the Socialization Subdomains. Factor III only contained the School/Community Subdomain and was not considered a meaningful factor. It was difficult to account for the cross loading evidenced by Interpersonal Relationships. It was unclear what part of the construct represented by this Subdomain is that which joined the other Factor I Subdomains, and what part of the construct is that which joined Factor II. All factor correlations were moderate to high (.590 to .737).

The correlated two-factor model demonstrated marginal fit with RMSEA = .083 and accounted for 65.7% of the variance. The Communication and Daily Living Skills Subdomains, and the Interpersonal Relationships Subdomain comprised Factor I. All three Socialization Subdomains comprised Factor II. Once again, Interpersonal Relationships demonstrated salient pattern coefficients on more than one factor. The correlation between Factors I and II (.662) was moderately high.

Finally, the one-factor model demonstrated poor statistical fit with RMSEA = .157 and accounted for 57.7% of the variance. Factor loadings indicated, with one exception (Personal, fair to good indicator), that the Subdomains were very good to excellent indicators of the factor.

Exploratory Bifactor Analysis

Based on the EFA results, a correlated two-factor model and the model that reflects the current Vineland-3 test structure were evaluated in the E-BFAs.

Interview form. Table 4 contains results for the two E-BFAs conducted on the Interview form.

[Insert Table 4 here]

The bifactor model which featured the General and two group factors evidenced good statistical fit with RMSEA = .060. The General factor accounted for 46.9% of the total variance and 79.1% of the common variance. The magnitude of factor loadings indicated that the Subdomains were good to excellent indicators of the General factor. Squaring the factor loadings indicated that the General factor accounted for 32.6% (Coping) to 71.7% (Written) of the variance in individual Subdomains. With respect to the group factors, Factor I Subdomains accounted for an additional 8.1% of the total variance and Factor II Subdomains an additional 4.3%. Collectively, the General and two group factors accounted for 59.3% of the total variance in Vineland-3 scores.

Table 4 shows the Omega coefficients. The GD_H for the General factor (.821) was high and indicated that it contained sufficient true score variance to be an interpretable measurement scale. For the group factors GD_{HS} for Factor I (.199) and Factor II (.121) were low which indicated they contained an insufficient amount of true score variance for interpretation as measurement scales. Table 4 also shows the E-BFA results for the model that represents the current Vineland-3 test structure: a General factor and three group factors. The model demonstrated good statistical fit with RMSEA = .005 and its ECVI (.176) was lower than that for the model with two group factors (.206). The General factor, which represents the ABC, accounted for 48.9% of the total variance and 75.7% of the common variance. The magnitude of factor loadings indicated that the Subdomains were good to excellent indicators of the General factor. The General factor accounted for 31.6% (Domestic) to 71.4% (Written) of the variance in the individual Subdomains. With respect to the group factors, which represent the Vineland-3 Domains, Communication accounted for an additional 4.3% of the total variance, Daily Living Skills an additional 5.8%, and Socialization an additional 5.7%. Together, the General and group factors accounted for 64.7% of the variance in Vineland-3 scores.

The GD_H coefficient for the General factor (.855) was large and indicated that it contained a sufficient amount of true score variance to justify interpretation of the ABC score. The GD_{HS} coefficient for the group factors were low (.089 to .241) and indicated that they did not possess a sufficient amount of true score variance to justify interpretation of the Domain scores.

Parent/Caregiver form. Table 5 presents the E-BFA results for the Parent/Caregiver form.

[Insert Table 5 here]

The bifactor model with the two group factors that emerged from the EFAs demonstrated good statistical fit with RMSEA = .068. The General factor accounted for 53.2% of the total variance and 81.5% of the common variance. The magnitude of factor loadings indicated that the Subdomains were very good to excellent indicators of the General factor. The General factor accounted for 40.6% (Domestic) to 66.3% (Written) of the variance in the individual Subdomains. With respect to the group factors, Factor I accounted for an additional 6.0% and

Factor II an additional 6.1% of the total variance. Together, the General and group factors accounted for 65.3% of the variance in the Vineland-3.

The pattern of omega coefficients was similar to that observed for the Interview form. The GD_H for the General factor (.871) was large and supported its interpretability as a measurement scale. However, GD_{HS} for Factor I (.073) and Factor II (.219) were low and did not support their interpretability as factor-based measurement scales.

Table 5 also shows the E-BFA results for the model that represents the current Vineland-3 test structure. This model demonstrated good statistical fit with RMSEA = .031 and its ECVI (.181) was lower than that for the two-group bifactor model (.218). The General factor accounted for 51.8% of the total variance and 80.1% of the common variance. The magnitude of Subdomain factor loadings indicated that they were good to excellent indicators of the General factor. The General factor accounted for 34.3% (Expressive) to 61.9% (Written) of the variance in the individual Subdomains. With respect to the group factors, which represent the Vineland-3 Domains, Communication accounted for an additional 6.2%, Daily Living Skills an additional 1.1%, and Socialization an additional 5.6% of the total variance. Together, the General and group factors accounted for 64.7% of the variance in Vineland-3 scores.

The G_{H} coefficient for the General factor (.871) was very high and supported the interpretation of the ABC score. The G_{HS} coefficients for the group factors were low (.033 to .208) and indicated that they contain too little true score variance to justify the interpretation of Domain scores.

Teacher form. Table 6 contains the E-BFA results for the Teacher form.

[Insert Table 6 here]

The bifactor model containing the two group factors that emerged from the EFAs evidenced good statistical fit with RMSEA = .066. The result should be interpreted with caution. The result indicated a Heywood case where 100% of the variance in the School/Community Subdomain was accounted for by the General and group factors. However, the result was consistent with E-BFA results from the Interview and Parent/Caregiver forms. The General factor accounted for 56.7% of the total variance and 82.6% of the common variance. The magnitude of factor loadings indicated that the Subdomains were good to excellent indicators of the General factor. The General factor accounted for 28.5% (Personal) to 85.0% (School/Community) of the variance in the individual Subdomains. Factor I accounted for an additional 3.5% and Factor II an additional 8.4% of the total variance. The General and two group factors collectively accounted for 68.6% of the total variance in Vineland-3 scores.

The G_H coefficient for the General factor (.872) was high and supported its interpretability as a measurement scale. However, the low G_{HS} values for Factor I (.061) and Factor II (.314) did not support their interpretability as factor-based measurement scales.

Table 6 shows that the bifactor model reflecting the current Vineland-3 test structure evidenced good statistical fit with RMSEA = .029. The ECVI (.173) was lower than that for the two-group bifactor model (.209). Once again, a Heywood case was observed with 100% of the variance in Numeric accounted for by the factors. While the result should be interpreted with caution, the pattern of results is consistent with those observed for the Interview and Parent/Caregiver forms. The General factor accounted for 56.3% of the total variance and 77.5% of the common variance. The magnitude of factor loadings indicated that the Subdomains were fair to excellent indicators of the General factor. The General factor accounted for 26.8% (Personal) to 87.8% (School/Community) of the variance in individual Subdomains. With respect to the group factors, Communication accounted for an additional 3.3% of the variance, Daily Living Skills 4.6%, and Socialization 8.4%. Together, the General and group factors accounted for 72.6% of the variance in Vineland-3 scores.

The GD_H coefficient for the General factor (.876) was very high and supported the interpretability of the ABC score. The GD_{HS} values for the group factors were very low (.101 to .310) and indicated that they possess too little true score variance to justify interpretation of the Domain scores.

Discussion

General Findings

We used factor analysis to examine the structural validity and Domain interpretability of the Vineland-3 Comprehensive forms for individuals in the 11 to 20-year age range because the technical manual did not report on these types of analyses. EFAs did not support the current three Domain test structure. Perhaps the most significant findings to emerge from this study were those from the E-BFAs. For all Vineland-3 forms, the pattern of results was remarkably consistent: (a) the General factor (ABC) accounted for far more of the total and common variance relative to the group factors (Domains); (b) the General factor (ABC), evidenced a high level of reliability and supported the clinical interpretation of the ABC score; and (c) the reliability for group factors (Domains), was low and indicated that they contained too little true score variance to justify their clinical interpretation. This means that the Domain scores do not add much unique reliable information about an individual's adaptive behavior performance beyond that which is already accounted for by the ABC score.

Perhaps this study's bifactor results are not surprising given the factor analytic results reported for the Vineland-2 Survey Forms (see Sparrow et al., 2005). Statistical fit indices provided mixed support for a higher-order three factor model and the reported factor loadings indicated the presence of a strong General factor (ABC). However, no data were presented to show the extent to which variance in Subdomain scores was accounted for by the General factor (ABC) versus the group factors (Domains). The current study's exploratory bifactor methods addressed these psychometric issues for the latest edition of the test, the Vineland-3.

Implications for Practitioners

The Vineland-3 technical manual provides information to guide practitioners on how they can use and interpret the ABC, Domain, and Subdomain scores but the present results do not empirically support some of the recommended practices. Results emphasize the importance of following the recommendations to practitioners by the AAIDD (2010), and others (e.g., XXXX, 2012; XXXX, 2006) to use a multimethod approach when evaluating individuals with developmental conditions. For example, the technical manual states a significant adaptive behavior deficit is indicated if the ABC or at least one Domain score falls two or more standard deviations below the population mean. Results of this study support the use of the ABC score for such decision-making because analyses indicated that most of the reliable variance in Domain scores is actually accounted for by a General factor (ABC). The ABC was found to be reliable and all nine Subdomains were generally strong indicators of that factor. Consistent with the Sparrow et al. (2016) definition of adaptive behavior in the context of the Vineland-3, the ABC score appears to reliably quantify an individual's performance of a broad range of activities required for personal and social self-sufficiency. Practitioners should use the ABC to determine an individual's overall level of adaptive performance and to use this score in conjunction with other data sources to substantiate diagnostic decisions.

Moreover, according to the Vineland-3 technical manual, a strengths and weaknesses analysis can be performed by examining statistically significant differences in adaptive performance among the Domain scores. The findings from the current study calls into question the usefulness of this analytic approach. The ABC provides the most reliable information about an individual's adaptive performance across the areas of communication, daily living, and socialization. Because the Vineland-3 purports to measure the *performance* of activities required for personal and social self-sufficiency and not adaptive *skills* per se, data obtained from other methods of direct assessment such as speech-language, academic, or occupational therapy testing would enable the practitioner to more reliably determine areas in need of intervention in one or more specific skill areas. Additional assessment should evaluate for other moderators of adaptive performance such as motivation deficits (e.g., through functional behavioral assessment, preference assessment) and the presence of co-occurring mental and physical health conditions. Such data should be collected across all relevant settings/contexts to inform intervention planning.

Finally, the Vineland-3 manual indicates the test can be used for progress monitoring by looking at changes in the ABC and Domain scores over time. Although our study did not examine the use of the Vineland-3 for progress monitoring per se, information from other tests can be used with the ABC as part of a progress monitoring protocol. For example, a practitioner may use the ABC score to evaluate the effect of targeted social communication training on overall adaptive behavior performance in an individual with autism spectrum disorder.

In sum, our findings support use of the Vineland-3, particularly the ABC score, within a multimethod assessment protocol, as recommended by the AAIDD (2010) and others (e.g., XXXX, 2012; XXXX, 2006), when assessing adolescents with developmental conditions. Information from multiple informants can be gathered on adaptive performance in different settings/contexts. Additional direct assessment methods can be used for the identification of potential moderators of adaptive performance such as developmental skill levels, motivation, and co-occurring conditions. All such information helps inform decisions related to support and intervention planning.

Limitations and Future Directions

One limitation of the present study pertains to the results for the Teacher form. The three correlated factors EFA, and bifactor analyses with two and three group factors produced Heywood cases. We found that 100% of the variance in the School/Community Subdomain was accounted for in the bifactor model with two group factors and in the EFA with three correlated factors. The Numeric Subdomain had 100% of its variance accounted for in the bifactor model that reflected the current organization of the Vineland-3. It is possible that these models reflected an over factoring of the data (see McDonald, 1985 for an overview). Whether one accepts a two-correlated factors model of the Teacher form or a statistically poor fitting one-factor model our results do not support the three Domain structure of the Vineland-3. Interestingly, the General factor (ABC) in the two and three group bifactor models accounted for 56.7% of the total variance, respectively. These percentages were comparable to the 57.7% of the total variance accounted for in the one-factor EFA model. Further, the General factor (ABC) in the two (82.6%) and three group (77.5%) bifactor models explained far more common variance than the group factors (Domains). Despite the loss in statistical fit it appears

that relatively little information about adaptive behavior is lost when accepting a one-factor model.

Replication is needed to evaluate the generality of the present results. Unfortunately, data were not available to replicate the results of the EFAs on an independent sample particularly through confirmatory factor analysis. Apart from the standardization sample, future factor analytic studies that include bifactor analysis of Vineland-3 data collected from specific subgroups within the general population, such as those with intellectual disability or autism spectrum disorder, would be highly informative. These are two subgroups for whom adaptive behavior assessment is critical. It is also noted that the present results were derived from the United States standardization sample and it is not known how they might apply to those in other countries.

Test developers for future editions of the Vineland should empirically evaluate the structural validity of the test to develop recommendations for its use. This includes applying methods such as bifactor analysis to determine the extent to which Domain scores provide unique and reliable information about an individual's adaptive performance that is not already accounted for by the overall ABC score. Item-level factor analysis can help determine the extent to which each Subdomain reflects a single narrow facet of adaptive performance. Finally, to complement the present study, the structural validity and Domain interpretability of the Vineland-3 Comprehensive forms should be evaluated for those age ranges not included in the present study. This would complete our understanding of the essential psychometric features of the Vineland-3 throughout the age ranges it covers and inform practitioners about its utility in a multimethod assessment protocol for individuals with developmental conditions.

22

References

- American Association on Intellectual and Developmental Disabilities (2010). Intellectual
 Disability: Definition, Classification, and Systems of Support. The AAIDD Ad Hoc Committee
 on Terminology and Classification, 11th Edition. American Association on Intellectual and
 Developmental Disabilities.
- American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders: 5th Edition*. Author.
- Bonifay, W.E., Reise, S.P., Scheines, R., & Meijer, R.R. (2015). *Structural Equation Modeling: A Multidisciplinary Journal*, 22(4), 504-516.
- Borthwick-Duffy, S.A. (2007). *Adaptive Behavior*. In, Handbook of Intellectual and Developmental Disabilities, Jacobson, Mulick, & Rojhan (Eds). Springer.
- Brown, T. A. (2006). Confirmatory factor analysis for applied research. Guilford Press.
- Browne, M.W. & Cudeck, R. (1989). Single sample cross-validation indices for covariance structures. *Multivariate Behavioral Research*, *24*, 445-455.
- Browne, M.W., Cudeck, R., Tateneni, K., & Mels, G. (2008). CEFA: Comprehensive Exploratory Factor Analysis, Version 3.01 [Computer software and manual]. http://faculty.psy.ohio-state.edu/browne/.
- Canivez, G.L., Watkins, M.W., & Dombrowski, S.C. (2016). Factor structure of the Wechsler Intelligence Scale for Children- Fifth Edition: Exploratory factor analyses with the 16 primary and secondary subtests. *Psychological Assessment*, 26(8), 975-986.
- Comrey, A.L. & Lee, H.B. (1992). *A first course in factor analysis (2nd ed.)*. Lawrence Erlbaum Associates.
- 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.

Individuals with Disabilities Education Improvement Act of 2004, Pub. L. 108-446., 118 Stat. 2647 (2006).

- XXXX. (2012) Considerations for establishing a multi-tiered problem- solving model for students with autism and emotional-behavioral disorders. *Psychology in the Schools, 49*: 975-987.
- XXXX. (2006) *Psychoeducational assessment*. In J.W. Jacobson, JA Mulick, Rojhan J. (Eds.),Handbook of Mental Retardation and Developmental Disabilities. Springer.

McDonald, R.P. (1985). Factor analysis and related methods. Lawrence Erlbaum Associates.

National Center for Education Statistics. (2014). *Children 3 to 21 years old served under Individuals with Disabilities Education Act (IDEA), Part B, by type of disability: Selected years, 1976-77 through 2012-13.*

Osborne, J.W. (2014). Best practices in exploratory factor analysis. Author.

- Pedhazur, E.J. & Schmelkin, L.P. (1991). *Measurement, design, and analysis: An integrated approach*. Lawrence Erlbaum Associates.
- Price, J.A., Morris, Z.A., & Costello, S. (2018). The application of adaptive behavior models: A systematic review. Behavioral Sciences, 8(1). https://doi.org/10.3390/bs8010011.
- Reise, S.P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, 47(5), 667-696.
- Reise, S.P., Moore, T.M., & Haviland, M.G. (2010). Bifactor models and rotations: Exploring the extent to which multidimensional data yield univocal scale scores. *Journal of Personality Assessment*, 92(6), 544-559.

Reise, S.P., Morizot, J., & Hays, R.D. (2007). The role of the bifactor model in resolving

dimensionality issues in health outcome measures. Quality of Life Research, 16, 19-31.

- Rodriguez, A., Reise, S.P., & Haviland, M.G. (2016). Applying bifactor statistical indices in the evaluation of psychological measures. *Journal of Personality Assessment*, *98*(*3*), 223-237.
- Sparrow, S.S., Cicchetti, D.V, & Saulnier, C.A. (2016). *Vineland Adaptive Behavior Scales, Third Edition* (Vineland-3). Pearson.
- Sparrow, S.S., Cicchetti, D.V, & Balla, D.A. (2005). Vineland Adaptive Behavior Scales, Second Edition (Vineland-2). AGS Publishing.
- Steiger, J.H., & Lind, J. (1980, May). *Statistically based tests for the number of common factors*. Paper presented at the annual meeting of the Psychometric Society, Iowa City.
- Tabachnick, B.G. & Fidell, L.S. (2007). Using multivariate statistics (5th ed.). Pearson Education, Inc.

Watkins, M. W. (2013). Omega [Computer software]. Ed & Psych Associates.

Table 1

			Pattern	Co	efficie	nts ^a	 			
Subdomain	Ι	u	I		II	u	Ι	II	III	u
Receptive	.719	.483	.5	66	.184	.495	 .681	.039	007	.491
Expressive	.728	.471	.6	70	.087	.460	.943	226	.044	.393
Written	.739	.454	.3	05	.498	.440	.686	.070	368	.263
Personal	.705	.504	.2	12	.568	.459	.177	.589	047	.447
Domestic	.644	.585	.2	15	.503	.544	020	.766	.098	.440
Community	.692	.521	(042	.858	.313	.051	.725	186	5.354
Interpersonal Rel.	.767	.412	.9	03	092	.297	.825	.002	.336	.254
Play & Leisure	.715	.489	.6	26	.127	.478	.725	.001	.048	.477
Coping	.694	.518	.7	11	.020	.474	.571	.172	.289	.441
Factor Correlation:			Ι		-		I	_		
			II	.7	22		II .8	01	-	
							III()91()84	-
RMSEA (90% CI)=	= .120	(.105, .13	5) .0	99	(.081,	.117)	.060	(.036,	.085)	
ECVI (90% CI)	.518	(.429, .62	3) .3	33	(.273,	.409)	.206	(.178,	.250)	

Exploratory Factor Analysis Results: Interview Form (N=480)

Note.

^aI = Factor 1; II = Factor 2; III= Factor 3; u = uniqueness.

Boldface indicates salient pattern coefficient or factor loading (\geq .32).

Table 2

Exploratory Factor Analysis Results: Parent/Caregiver Form (N=480)

Pattern Coefficients ^a										
Subdomain	Ι	u	Ι	Π	u]	I	II	III	u
Receptive	.700	.510	.788	057	.442		.494	024	.440	.418
Expressive	.681	.536	.536	.192	.526	-	001	.251	.810	.114
Written	.774	.400	.838	026	.329		.729	030	.221	.330
Personal	.794	.369	.787	.042	.332		.660	.049	.227	.339
Domestic	.644	.585	.409	.276	.593		.520	.238	071	.537
Community	.722	.479	.592	.163	.483		.732	.099	055	.385
Interpersonal Rel.	.773	.402	020	.937	.148	-	014	.912	.029	.163
Play & Leisure	.749	.439	.334	.477	.430		.234	.484	.172	.424
Coping	.758	.426	.237	.600	.378		.273	.591	018	.361
Factor Correlations:										
			I	-			Ι			
			II .72	24			II	.701		
							III	.391	.409	
RMSEA(90% CI)=	.141	(.127, .156)	.121	(.104,	.139)		.066 ((.043,	.090)	
ECVI (90% CI) =	.671	(.566, .793)	.427	(.352,	.517)		.209 ((.179,	.254)	

Note.

 ${}^{a}I =$ Factor 1; II = Factor 2; III- Factor 3; u = uniqueness. **Boldface** indicates salient pattern coefficient or factor loading (\geq .32).

Table 3

			Pattern (Coeffi	cients ^a				
Subdomain	Ι	u	Ι	II	u	Ι	II	III	u
Receptive	.869	.245	.814	.086	.237	.658	.050	.226	.241
Expressive	.805	.353	.768	.067	.338	.821	.112	070	.294
Written	.849	.280	.914	041	.212	.751	047	.200	.227
Personal	.559	.687	.380	.229	.688	.439	.261	068	.665
Numeric	.736	.458	.838	094	.392	.781	058	3.044	.391
School/Community	.855	.269	.617	.308	.272	.043	.048	.936	.000
Interpersonal Rel.	.792	.373	.369	.540	.309	.367	.521	.054	.300
Play & Leisure	.634	.598	030	.843	.322	.018	.836	009	.293
Coping	.674	.546	.051	.790	.320	008	.683	.192	.339
Factor Correlation:			I	-		I			
			II .6	62	-	II .5	90	-	
						III .7	37.6	31	-
RMSEA (90% CI)=	.157	(.142, .171)	.083	(.066,	.102)	.068	(.045,	.093)	
ECVI (90% CI) =	.788	(.673, .917)	.274	(.224,	.339)	.219	(.187,	.266)	

Exploratory Factor Analysis Results: Teacher Form (N=500)

Note.

^aI = Factor 1; II = Factor 2; III = Factor 3; u = uniqueness. **Boldface** indicates salient pattern coefficient or factor loading (\geq .32).

Table 4

				Bifacto	or Loadings ^a					
							Daily			
Subdomain	G	Ι	II	u	G	Communication	Living Skills	Socialization	u	
Receptive	.677	.225	01	0.491	.769	.594	030	014	.055	
Expressive	.714	.259	17	3.393	.738	.019	102	.126	.428	
Written	.847	082	211	1.263	.845	173	011	182	.222	
Personal	.682	.068	.288	.447	.652	.035	.359	.039	.443	
Domestic	.591	.142	.437	.440	.562	.023	.464	.162	.442	
Community	.733	083	3.320	.354	.678	099	.418	067	.352	
Interpersonal Rel.	.648	.566	.075	.254	.714	017	.016	.592	.139	
Play & Leisure	.665	.284	01	5.477	.687	057	.018	.203	.484	
Coping	.571	.455	.161	.441	.609	.136	.138	.348	.471	
RMSEA (90% CI)=	.060	(.036,	.085))	.005	(.000, .060)				
ECVI (90% CI)=	.206	(.178,	.250))	.176	(.175, .197)				
GD^{b}	.920	.868	.860		.930	.889	.800	.825		
$\mathrm{GO}_{\mathrm{H}}/\mathrm{GO}_{\mathrm{HS}}^{\mathrm{c}}$.821	.199	.121		.855	.089	.241	.202		
Total Variance	.469	.081	.043		.489	.043	.058	.057		
ECV ^d	.791	.136	.073		.757	.066	.089	.088		

Bifactor Results For Vineland-3 Interview Form (N=480)

Note. ^aG = General Factor; I= Factor I; II= Factor 2; u = uniqueness.

^bOmega coefficient GD= $[(\sum \lambda_{Gen})^2 + (\sum \lambda_{Grp1})^2 + ... (\sum \lambda_{GrpP})^2]/[(\sum \lambda_{Gen})^2 + (\sum \lambda_{Grp1})^2 + ... (\sum \lambda_{GrpP})^2 + \sum (1-h^2)].$ ^cOmega hierarchical GD_H= $[(\sum \lambda_{Gen})^2]/[(\sum \lambda_{Gen})^2 + (\sum \lambda_{Grp1})^2 + ... (\sum \lambda_{GrpP})^2 + \sum (1-h^2)]$ /Omega hierarchical subscale $GD_{HS} = [(\sum \lambda_{GrpP})^2]/[(\sum \lambda_{GrpP})^2 + \sum (1-h^2)].$

^dExplained common variance.

Italics indicate: (a) group factor assignments based on EFA two-factor results and (b) Vineland-3 test structure. **Boldface** indicates salient factor loading (\geq .32).

Bifactor Results For Vineland-3 Parent/Caregiver Form (N=480)

					Bifactor Loa	ading	s ^a			
						U		Daily		
Subdomain	G	Ι	II	u		G	Communication	Living Skills	Socializa	ation u
Receptive	.729	.205	090	.418		.712	.402	132	136	.296
Expressive	.697	.627	.077	.114		.586	.607	.024	.204	.246
Written	.814	029	084	.330		.787	.154	.129	074	.335
Personal	.812	006	030	.339		.781	.158	.106	001	.353
Domestic	.637	199	.131	.537		.683	133	064	.050	.510
Community	.745	245	.028	.385		.771	149	.294	.004	.296
Interpersonal Rel.	.705	.024	.583	.163		.721	.040	082	.508	.214
Play & Leisure	.702	.067	.279	.424		.666	.145	.228	.383	.337
Coping	.705	091	.366	.361		.749	058	153	.314	.313
RMSEA (90% CI)=	.068 ((.045,	.093)			.031	(.000, .072)			
ECVI (90% CI)=	.219 ((.187,	.266)			.181	(.175, .207)			
$\ensuremath{\mathrm{GD}}^{\mathrm{b}}$.936	.908	.861			.934	.854	.809	.861	
GD_{H}/GD_{HS}^{c}	.871	.073	.218			.871	.203	.033	.208	
Total Variance	.532	.060	.061			.518	.062	.011	.056	
ECV ^d	.815	.091	.094			.801	.095	.017	.086	

Note. ${}^{a}G$ = General Factor; u = uniqueness.

^bOmega coefficient GD= $[(\sum \lambda_{Gen})^2 + (\sum \lambda_{Grp1})^2 + ...(\sum \lambda_{GrpP})^2]/[(\sum \lambda_{Gen})^2 + (\sum \lambda_{Grp1})^2 + ...(\sum \lambda_{GrpP})^2 + \sum(1-h^2)].$ ^cOmega hierarchical GD_H= $[(\sum \lambda_{Gen})^2]/[(\sum \lambda_{Gen})^2 + (\sum \lambda_{Grp1})^2 + ...(\sum \lambda_{GrpP})^2 + \sum(1-h^2)]$ Omega hierarchical subscale $GD_{HS} = [(\sum \lambda_{GrpP})^2]/[(\sum \lambda_{Gen})^2 + (\sum \lambda_{GrpP})^2 + \sum(1-h^2)].$

^dExplained common variance.

Italics indicate: (a) group factor assignments based on EFA two-factor results and (b) Vineland-3 test structure. **Boldface** indicates salient pattern coefficient or factor loading (\geq .32).

Table 6

Bifactor Results For Vineland-3 Teacher Form (N=500)

Bifactor Loadings ^a									
Subdomain	G	Ι	II	u	G	Communication	Daily Living Skills	Socialization	u
Receptive	.865 .0	98	005	.241	.840	.219	012	.010	.247
Expressive	.795 .2	70	.024	.294	.754	.464	.046	.051	.212
Written	.864 .1	38	084	.227	.839	.181	.076	053	.254
Personal	.534 .1	55	.160	.665	.518	.171	.090	.176	.664
Numeric	.744 .2	13	097	.391	.806	037	.577	129	.000
School/Community	.922	384	.048	.000	.937	139	275	.011	.026
Interpersonal Rel.	.752 .0	79 .	.358	.300	.736	.168	034	.362	.298
Play & Leisure	.575(002 .	.613	.293	.582	030	014	.629	.265
Coping	.629 (093 .	.506	.339	.644	053	084	.478	.347
RMSEA (90% CI)= ECVI (90% CI)=	.066 (.04 .209 (.1	43, . 79, .	090) 254)		.029 (.173 (.000, .070) .168, .197)			
GD ^b GD _H /GD _{HS} ^c Total Variance ECV ^d	.945 .9 .872 .0 .567 .0 .826 .0	28 61 35 51	.864 .314 .084 .123		.951 .876 .563 .775	.902 .101 .033 .045	.884 .131 .046 .064	.863 .310 .084 .116	

Note. ^aG = General Factor; I= Factor 1; II= Factor 2; u = uniqueness.

^bOmega coefficient GD= $[(\sum \lambda_{Gen})^2 + (\sum \lambda_{Grp1})^2 + ... (\sum \lambda_{GrpP})^2]/[(\sum \lambda_{Gen})^2 + (\sum \lambda_{Grp1})^2 + ... (\sum \lambda_{GrpP})^2 + \sum (1-h^2)]$ ^cOmega hierarchical GD_H= $[(\sum \lambda_{Gen})^2]/[(\sum \lambda_{Gen})^2 + (\sum \lambda_{Grp1})^2 + ... (\sum \lambda_{GrpP})^2 + \sum (1-h^2)]$ /Omega hierarchical subscale $GD_{HS} = [(\sum \lambda_{GrpP})^2]/[(\sum \lambda_{Gen})^2 + (\sum \lambda_{GrpP})^2 + \sum (1-h^2)].$

^dExplained common variance.

Italics indicate: (a) group factor assignments based on EFA two-factor results and (b) Vineland-3 test structure. **Boldface** indicates salient pattern coefficient or factor loading (\geq .32).



Figure 1. Bifactor model of the Vineland-3 Interview/Parent Caregiver Form. General Factor: ABC= Adaptive Behavior Composite; Group Factors: COM= Communication Domain; DLS= Daily Living Skills Domain; SOC= Socialization Domain; Subdomains: REC= Receptive; EXP= Expressive; WRI= Written; PER= Personal; DOM= Domestic; COM= Community; IPR= Interpersonal Relationships; PLS= Play and Leisure Skills; COP= Coping.