HHS Public Access

Author manuscript

J Intellect Disabil Res. Author manuscript; available in PMC 2020 January 03.

Published in final edited form as:

J Intellect Disabil Res. 2020 January ; 64(1): 18–26. doi:10.1111/jir.12691.

Concordance of the Vineland Adaptive Behavior Scales, Second and Third Editions

Cristan Farmer[1], **Dee Adedipe**[1], **Vanessa Bal**[2], **Colby Chlebowski**[1], **Audrey Thurm**[1]

[1]Neurodevelopmental and Behavioral Phenotyping Service, Intramural Research Program, National Institute of Mental Health

[2]Graduate School of Applied and Professional Psychology, Rutgers University

Abstract

Background: Due to its centrality in the conceptualization of intellectual disability, reliable and valid measurement of adaptive behaviour is important to both research and clinical practice. The manual of the Vineland Adaptive Behavior Scales, recently released in its third edition, provides limited reliability information obtained from a sample composed primarily of typically developing individuals. The goal of this study was to evaluate the concordance of the Vineland-3 with the Vineland-II in a sample more similar in ability level to those in which the Vineland is commonly used.

Methods: Both editions of the Vineland Comprehensive Parent Interviews were conducted with a convenience sample of 106 parents/caregivers of individuals with neurodevelopmental disability, participating at two neurodevelopmental disorder research clinics. Administrations were up to 7 days apart, but most (90%) were simultaneous. The concordance correlation coefficient (CCC) [95% CI] and mean differences [95% CI] were calculated for domain standard scores and subdomain *v*-scale scores.

Results: Domain-level CCC ranged from 0.78 [0.70, 0.84] for Communication to 0.86 [0.76, 0.92] for Motor. Subdomain CCC ranged from 0.71 [0.62, 0.78] for Receptive Language to 0.91 [0.85, 0.95] for Fine Motor. Vineland-3 scores were lower than Vineland-II scores; 77% of participants had lower Adaptive Behavior Composite scores on the Vineland-3 than on the Vineland-II. For the subdomains, the magnitude of this difference depended upon the level of adaptive behaviour. For Communication, the domain with the lowest CCC, the mean difference ranged from −13.70 [−8.03, −19.35] for a Vineland-II score or 85 to a difference of −19.18 [−12.28, −26.87] for a Vineland-II score of 40.

Discussion: Among individuals with intellectual and developmental disabilities, the Vineland-3 produces lower scores than the Vineland-II, and these clinically significant differences tend to be larger for individuals with lower levels of ability. Thus, care must be taken in interpreting scores from the Vineland-3 relative to those obtained from the previous edition.

Correspondence: Cristan Farmer, 10 Center Drive, Bethesda MD 20892. Ph: (301) 435-3999.

Conflicts of interest: CF, DA, VB, CC, and AT report no real or potential conflicts of interest.

Keywords

Adaptive behaviour; intellectual disability; developmental disorder; psychometrics; reliability; concordance

INTRODUCTION

Adaptive behaviour is used to operationalise intellectual disability (ID); diagnostic criteria require that both cognitive ability and adaptive behaviour are significantly impaired during childhood or adolescence (American Psychiatric Association, 2013, World Health Organization, 2018). Because adaptive behaviour refers broadly to skills in conceptual, social, and practical domains, individuals who meet diagnostic criteria, which require functional impairment, for any neurodevelopmental disability are likely to exhibit deficits in adaptive behaviour. As described by Tasse and colleagues (2012), “adaptive behavior encompasses an essential dimension in a multidimensional understanding of human functioning,” (p. 291) and is therefore routinely used in cross-sectional and longitudinal studies of neurodevelopmental disability.

Due to its centrality in the ID diagnostic criteria, several standardised instruments of adaptive behaviour exist. Most available instruments are parent- or caregiver-interview or report [e.g., Scales of Independent Behavior, Revised (Bruininks et al., 1996), Adaptive Behavior Scale – School edition (Lambert et al., 1993)], although some self-report instruments do exist (e.g., the Adaptive Behavior Assessment System, 3rd edition has both other and self-report forms (Harrison and Oakland, 2015)). Here, we focus on the Vineland Adaptive Behavior Scales Interview Form, which was recently published in its third edition (Sparrow et al., 2016).

Several important changes from the Vineland-II are reflected in the Vineland-3. The overall number of items on the scale increased by 34% relative to the Vineland-II (see Supplementary Materials). Many of the new items reflect earlier/easier skills and are therefore at lower levels of ability. These new items are concentrated in a few subdomains (i.e., Receptive Communication, Written Communication, Personal Skills, and Community Skills) (Saulnier, 2016). The Vineland-3 manual also gives new instructions for assigning a score of 1. In the Vineland-II a score of 1 was assigned when a behaviour was “sometimes” or “partially” observed; in the Vineland-3, a score of 1 is assigned when the behaviour is observed “sometimes.” A Vineland-II rating of “partially” allowed for the possibility that prompting of the behaviour may have occurred; in the Vineland-3, the behaviour must be observed to occur spontaneously (although exceptions to this for a variety of items are described in the manual). The start age for some subdomains were adjusted to address the concern that some items on the Coping, Domestic, and Community domains were inappropriate for young children. Finally, while the Vineland-II included the Motor Skills domain in the computation of the Adaptive Behavior Composite for children younger than 7 years, this is not true for individuals of any age on the Vineland-3. Other more minor changes such as relocation of items between domains and updates to the language are described in the manual.

As with any new edition, it is important to consider how content, administration, and scoring differences may affect the score of an individual on the measure. This is particularly important for the Vineland, given that its use as an outcome measure in longitudinal studies and treatment trials is increasing (Farmer et al., 2018, Bal et al., 2015, Szatmari et al., 2015, Chatham et al., 2018, Farmer et al., 2019). The Vineland-3 manual does provide some information on the correspondence between the interview editions, administered to the same individual between 12 and 35 days apart (Sparrow et al., 2016). The mean difference (in standard deviation units) and the Pearson correlation between the Vineland-II and Vineland-3 scores were calculated for samples of children and young adults grouped by age (0–2, 3–6, 7–11, 12–20 years). Sample sizes ranged from 43 to 73.

Correlations between the editions were moderate, ranging from 0.40 (Communication in the oldest subgroup) to 0.87 (Communication, Socialization, and the Adaptive Behavior Composite in the 3 – 6 years subgroup). However, two scores may be strongly correlated even in the presence of a mean difference. Indeed, mean differences were nearly uniformly in the direction of lower Vineland-3 than Vineland-II scores, although Cohen's d effect sizes ranged from −0.62 (Receptive Language for the Age 3 – 6 group) to +0.61 (Receptive Language for the Age 12 – 20 group). There was no consistent pattern across the age groups, except that the mean differences were smallest for the youngest group (0 – 2 years), and mean differences were more likely to be positive (i.e., Vineland-3 scores higher than Vineland-II) for the 12 – 20 age group. The manual attributes the lower Vineland-3 scores possibly to content changes, but possibly to "more lenient …reporting" (p.222) by parents, caregivers, and teachers, although it is unclear why this would be expected in the short test-retest timeframe.

Importantly, although the level of functioning of the test-retest samples is not thoroughly described, the mean scores are very close to 100 and the sample standard deviation usually slightly less than 15, suggesting that very few participants were in the range of intellectual disability. Although a population representative sample is essential for the development of normative data, it is possible that the psychometric performance of the Vineland in such a sample is not relevant to its performance in a sample with neurodevelopmental disabilities.

For both clinical and research applications, it is necessary to understand the difference in scores yielded by the second and third editions of the Vineland in the populations for which the Vineland is most frequently used. The goal of this study was to evaluate the concordance of the Vineland-3 with the Vineland-II, with special attention to use with individuals with lower levels of ability. We hypothesised that the correlation between scores from the two tests would be very strong, but that concordance estimates would be more moderate, reflecting our hypothesis that Vineland-3 scores would be lower than Vineland-II scores on the same participants.

METHODS

Participants

Data were prospectively collected during 2017 and 2018 from a sample of convenience at two neurodevelopmental disorder research clinics (see demographic information in Table 1).

Participants at both clinics were diagnosed with a variety of neurodevelopmental disorders, including autism spectrum disorder and/or various neurogenetic conditions including Down syndrome, Smith-Lemli-Opitz Syndrome, Williams syndrome, metabolic disorders, and other undiagnosed conditions which affect neurodevelopment early in life. The Neurodevelopmental and Behavioral Phenotyping Service in the Intramural Research Program of the National Institute of Mental Health is a research clinic. Both versions of the Vineland interview were administered to the parents/caregivers of 87 individuals enrolled in institutional review board approved research protocols sponsored by NIMH, NICHD and NHGRI at the NIH Clinical Center. The University of California San Francisco Service, Training, Advocacy and Research (UCSF STAR) Center is a multidisciplinary specialty neurodevelopmental disorders clinic. Participants were clinical (n=10) and research referrals (n=9) seen at STAR between 2017 and 2018.

Procedure

This study was approved by the institutional review boards at both sites. Depending on age and developmental level, participants or parents/legal guardians consented to the use of these data.

At both sites, the Vinelands were administered by a licensed clinical psychologist or by research assistants trained and supervised by a licensed clinical psychologist who had been trained on both the Vineland-II and Vineland-3 using materials provided by the publisher (i.e., manual, webinars) and trainings that included multiple coders for establishing initial reliability. Due to logistical constraints, it was not possible to ensure that both versions of the Vineland were administered by the same clinician. Of 106 cases, 63 were assessed by the same clinician for both Vineland versions. For this project, an effort was made to obtain both the Vineland-II and the Vineland-3 on the same day; administrations more than 7 days apart were excluded from the study. Simultaneous administrations were obtained for n=80 (91%) at NIH and for n = 15 (79%) at STAR (see Table 1).

The comprehensive interview versions of the Vineland-II and Vineland-3 were used. Both forms yield scores for four domains (Communication, Daily Living Skills, Socialization, and Motor Skills), up to 11 subscales (three each for all domains except Motor, which has two), and an overall Adaptive Behavior Composite Score. Standard scores (population mean = 100, SD = 15) are produced for the domain and composite scores, and scaled scores (called *v*-scale scores, with a population mean = 15, SD = 3) are produced for the subscale scores. We do not report on the maladaptive behaviour scales. Both forms are appropriate for use in individuals aged birth to 90 years, but some subscales are not administered to all ages. The Vineland-II has two age restrictions; one for the Written subscale (≥3 years) and one for the Motor Skills domain (≤ 7 years). For the Vineland-3, the Written, Domestic, and Community subscales have a minimum age of 3, the Coping Skills subscale has a minimum age of 2, and the Motor domain has a maximum age of 9 years, so standard scores could not be obtained for these domains in the specified exclusionary age ranges. While both versions of the Vineland provide suggested qualitative descriptors of “high” (domain and ABC Standard Scores of 130–140), “moderately high” (domain and ABC Standard Scores of 115–129), “adequate” (domain and ABC Standard Scores of 86–114), “moderately low” (domain

and ABC Standard Scores of 71–85), and "low" (domain and ABC Standard Scores of 20–70), these descriptors were not used due to insufficient precision at lower levels.

Statistical Analysis

The hypothesis in this study is about the concordance of the two versions of the Vineland, rather than their correlation, which may be high even in the presence of the hypothesised mean difference. The concordance correlation coefficient (CCC) is equivalent to the appropriately specified ICC (Shrout and Fleiss, 1979), and is expressed in terms of the variance components of a linear mixed model (Carrasco and Jover, 2003). Values closer to 1.0 reflect exact agreement between tests, and a threshold of 0.50 is often used as the lower bound of the "acceptable" range. Age and site were added as covariates. The cccrm package (Carrasco and Martinez, 2015) for R version 3.5 was used to estimate the CCC and the 95% confidence interval.

To calculate the estimated mean difference between the forms, a general linear model was used to predict Vineland-3 scores from Vineland-II scores, controlling for site and age. This form of the model was selected because it allows estimation of a mean difference that might differ depending on the ability level (i.e., both a slope and an intercept are estimated). In all analyses, age was centered at 5 years to ensure interpretable parameters for all scales and subscales. To yield an interpretable intercept, the standard scores were centered at the population mean (100 for standard scores and 15 for *v*-scale scores). For the purposes of interpretation, the mean difference was estimated at relevant levels of Vineland-II score (i.e., 85, 70, 55, and 40 for standard scores, and 15, 12, 9, and 6 for *v*-scale scores). Model assumptions were evaluated via visual inspection of the residuals. All data visualization was performed in SAS/STAT 9.4. All R and SAS syntax is provided in the Supplementary Materials.

As described above, several Vineland subscales have age restrictions, which create missing data when the individual is outside the designated age range. Further, due to administration error, scores on the Written Communication subscale (n=1) and the Adaptive Behavior Composite (n = 1) could not be computed. Listwise deletion was used per subscale; thus, the sample size for each ICC or general linear model depended on the domain/subscale: Adaptive Behavior Composite, n=105; Written Communication, n=93; Community, n=94; Domestic, n=94; Coping, n=101; Motor Skills (and both subdomains), n=44; all other domains/subscales, n=106.

No a priori power calculation was done, as this was a sample of convenience. However, based on published estimates (Zou, 2012), the sample size in this study is near to that required to demonstrate with 80% assurance that the confidence interval surrounding an ICC of 0.70 will exclude 0.50 (N = 101), which is often used as the lower limit of "acceptable." We adhere to the current American Statistical Association guidelines and emphasise the 95% confidence intervals for all hypothesis tests alongside the exact p-values, and do not refer to "statistical significance." (Wasserstein et al., 2019).

RESULTS

Participant demographics are described in Table 1. For each of the domains and subscales, the CCC estimates exceeded 0.70 (see Figure 1 and Table S1). There were not meaningful differences in concordance across the subscales, as evidenced by overlapping confidence intervals.

As hypothesised, Vineland-3 scores were lower than Vineland-II scores (Figure 2). At the domain level, differences between editions were largest for the Communication subscale and smallest for the Motor subscale. In regressions predicting Vineland-3 scores, Vineland-II slopes different from 1.0 indicate that the mean difference between editions depends on the adaptive behavior level of the examinee. For example, the estimated difference for a Vineland-II Communication score of 85 was about 14 points (estimated Vineland-3 score 71.31, 95% CI = 65.65 – 76.97), while the difference was about 17 for a Vineland-2 score of 55 (estimated Vineland-3 score 37.65, 95% CI = 31.57 – 43.74) (see Figure 2 and Tables S2 and S3).

To assist in interpretation of these discrepancies at the individual level, the axes of Figure 3 are categorised into Average or above (≥ 85), Borderline (70 – 84), Mild (55 – 69), Moderate (40 – 54), Severe (25 – 39), Profound (20 – 24) impairment levels. For the Adaptive Behavior Composite, six participants were classified by the Vineland-3 two levels lower than by the Vineland-II (e.g., Severe versus Mild).

DISCUSSION

This is the first independent study of the concordance between the Vineland-II and Vineland-3. While moderate concordance was observed, the Vineland-3 produces lower scores than the Vineland-II across all domains and ability levels. These findings are consistent with those published in the Vineland-3 manual, but they extend that work by evaluating the scales in a sample of individuals with various neurodevelopmental disabilities, rather than the population-based samples used in the manual. Thus, these psychometric data are more readily generalizable to the types of samples in which the Vineland is often used.

Even scores with a large mean difference may exhibit a strong correlation; for this reason it is important to characterise concordance using an agreement statistic rather than a Pearson correlation (Cicchetti, 2017). In this study, we chose the concordance correlation coefficient (CCC), which in this case is identical to the intraclass correlation calculated for absolute agreement. The lower boundaries of the CCC confidence intervals were generally around 0.70; while this is acceptable, it reflects the fact that scores from the two instruments do differ. Specifically, we found that the Vineland-3 produces scores that are lower than the Vineland-II, and this difference is larger at lower levels of adaptive behaviour. In some cases, the estimated mean difference in scores exceeded 15 points. Differences of this magnitude resulted in a clinically significant shift of ID classification to two levels below their Vineland-II classification.

The mean difference in scores between the Vineland-II and the Vineland-3 has important implications for clinical and research practice. Clinically, the Vineland is used to monitor an

individual's progress or to establish or confirm ongoing eligibility for support. Our results indicate that when the clinician adopts the Vineland-3, the adaptive behaviour estimates for a given client will be lower than they would have been with the Vineland-II. This decrease must be interpreted with caution, as it likely reflects a change in measure rather than a change in adaptive behaviour functioning. Similar caution is needed in research settings, where findings with the Vineland-3 will be compared to existing work using the Vineland-II. The difference in Vineland-3 and Vineland-II scores is especially relevant to treatment studies, as a change in measure could be mistaken for a treatment effect. Thus, these results provide a strong argument against transitioning between editions mid-study; should this occur, researchers must account for the form version in any statistical analysis. For the Adaptive Behavior Composite, the mean differences were relatively constant across levels of adaptive behaviour, so a main effect of form could sufficiently account for the edition change. However, particularly for the Communication domain, the mean difference in editions varied by level of function level. In this case, a simple solution is not evident.

Limitations

This was a sample of convenience, and we were therefore unable to adequately evaluate potential moderators of the reliability between forms, such as age. For example, multiple reports suggest that Vineland-II scores tend to decline with age in several neurodevelopmental disorders (Kanne et al., 2011, Klaiman et al., 2014, Ghezzo et al., 2014). This effect of age may moderate the reliability between Vineland-II and Vineland-3. Because of the relatively small sample and inadequate representation at all combinations of age and adaptive behaviour, we were unable to properly explore this possibility. Future work, especially that by the test developers, should address this possibility.

Other limitations stem from the fact that this was a sample of convenience, including that age restrictions on the Motor domain meant that those analyses were underpowered, and that the results of this study may not be generalizable to all clinical and research settings. Because the mean difference between editions depended upon adaptive behaviour level for some subscales, the composition of the sample directly affected the mean differences we obtained. To mitigate this fact, we estimated the mean differences at several reference levels. We did not randomly assign either the order in which the Vineland-II and Vineland-3 were presented or the clinician, and therefore cannot statistically account for these effects. The short timeframe between administrations, which were frequently simultaneous, makes this a true test of reliability, untainted by true changes in behaviour. In many cases, these were not truly independent administrations; for this reason, our estimates of reliability should be viewed as an upper limit on true reliability.

Finally, although concordance was the focus of this study, a question of validity arises. Which is a truer reflection of the adaptive function of the individual, the higher Vineland-II score or the lower Vineland-3 score? Unfortunately, there is no external gold standard for evaluation of adaptive behaviour, so we are unable to judge the editions' respective validity. Future work on this specific topic may include item-level content analyses to determine the extent to which the Vineland-3 domains differ from the Vineland-2. More broadly, however, the issue of the construct validity of adaptive behaviour is of renewed interest, as it is used to

characterize a range of conditions affecting neurodevelopment (Slomine et al., 2018, Delaney et al., 2014, Vago et al., 2011). The alignment of adaptive behaviour with the activity/participation and other elements of the WHO International Classification of Functioning may further bolster its adoption in areas other than intellectual disability (Gleason and Coster, 2012).

Conclusion

The recently published Vineland-3 Adaptive Behavior Scales (Comprehensive Interview Form) contains many changes, including new items and different scoring instructions. As with any standardised measure, it is important for both researchers and clinicians to understand the correspondence between the new and previous editions. In this study, we generalise the findings published in the Vineland-3 manual, which documented slightly lower Vineland-3 scores among typically developing individuals. Among individuals with intellectual and developmental disabilities, the Vineland-3 produces lower scores than the Vineland-II, and these clinically significant differences tend to be larger for individuals with lower levels of adaptive behaviour. Thus, care must be taken in interpreting scores from the Vineland-3 relative to those obtained from the previous edition.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments:

This work was supported in part by the Intramural Research Program of the National Institute of Mental Health (1ZICMH002961;) and K23MH115166 to VB. We thank Glennis Muldoon, Elizabeth Mistur, Bethany Sauls, Jerry D. Harrison, Megan Fok and Alexis Sullivan for their assistance in data collection.

REFERENCES

AMERICAN PSYCHIATRIC ASSOCIATION 2013 Diagnostic and statistical manual of mental disorders : DSM-5, Washington, D.C., American Psychiatric Association.

BAL VH, KIM S-H, CHEONG D & LORD C 2015 Daily living skills in individuals with autism spectrum disorder from 2 to 21 years of age. Autism, 19, 774–784. [PubMed: 25922445]

BRUININKS RH, WOODCOCK RWBK, WEATHERMAN RF & HILL BK 1996 Scales of Independent Behavior-Revised. SIB-R, Riverside Publishing Company Itasca, IL.

CARRASCO JL & JOVER L 2003 Estimating the generalized concordance correlation coefficient through variance components. Biometrics, 59, 849–858. [PubMed: 14969463]

CARRASCO JL & MARTINEZ JP 2015 Concordance Correlation Coefficient for Repeated (and Non-Repeated) Measures. R package version 1.2.1.

CHATHAM C, TAYLOR K, CHARMAN T, LIOGIER D'ARDHUY X, EULE E, FEDELE A, HARDAN A, LOTH E, MURTAGH L & DEL VALLE RUBIDO M 2018 Adaptive behavior in autism: Minimal clinically important differences on the Vineland-II. Autism Research, 11, 270–283. [PubMed: 28941213]

CICCHETTI D 2017 Choice of agreement statistics: A discussion of the underlying biostatistics, with heuristic examples from the Vineland-3 [Online]. NCS Pearson, Inc Available: https://images.pearsonclinical.com/images/assets/vineland-3/Vineland3_Biostatistics-Agreement_1A.pdf [Accessed].

DELANEY KA, RUDSER KR, YUND BD, WHITLEY CB, HASLETT PA & SHAPIRO EG 2014 Methods of neurodevelopmental assessment in children with neurodegenerative disease: Sanfilippo syndrome. JIMD Rep, 13, 129–37. [PubMed: 24190801]

FARMER C, SWINEFORD L, SWEDO SE & THURM A 2018 Classifying and characterizing the development of adaptive behavior in a naturalistic longitudinal study of young children with autism. Journal of neurodevelopmental disorders, 10, 1. [PubMed: 29329511]

FARMER CA, THURM A, FARHAT N, BIANCONI S, KEENER LA & PORTER FD 2019 Long-Term Neuropsychological Outcomes from an Open-Label Phase I/IIa Trial of 2-Hydroxypropyl-β-Cyclodextrins (VTS-270) in Niemann-Pick Disease, Type C1. CNS drugs, 1–7. [PubMed: 30511350]

GHEZZO A, SALVIOLI S, SOLIMANDO MC, PALMIERI A, CHIOSTERGI C, SCURTI M, LOMARTIRE L, BEDETTI F, COCCHI G & FOLLO D 2014 Age-related changes of adaptive and neuropsychological features in persons with Down Syndrome. PloS one, 9, e113111. [PubMed: 25419980]

GLEASON K & COSTER W 2012 An ICF-CY-based content analysis of the Vineland Adaptive Behavior Scales-II. J Intellect Dev Disabil, 37, 285–93. [PubMed: 22989139]

HARRISON PL & OAKLAND T 2015 ABAS-3: Adaptive behavior assessment system, Western Psychological Services Los Angeles, CA.

KANNE SM, GERBER AJ, QUIRMBACH LM, SPARROW SS, CICCHETTI DV & SAULNIER CA 2011 The Role of Adaptive Behavior in Autism Spectrum Disorders: Implications for Functional Outcome. Journal of Autism and Developmental Disorders, 41, 1007–1018. [PubMed: 21042872]

KLAIMAN C, QUINTIN EM, JO B, LIGHTBODY AA, HAZLETT HC, PIVEN J, HALL SS, CHROMIK LC & REISS AL 2014 Longitudinal profiles of adaptive behavior in fragile X syndrome. Pediatrics, 134, 315–24. [PubMed: 25070318]

LAMBERT N, NIHIRA K & LELAND H 1993 ABS-S: 2: AAMR Adaptive Behavior Scale: School, Pro-ed.

SAULNIER C 2016 Vineland Adaptive Behavior Scales, Third Edition [Slides] [Online]. http://downloads.pearsonclinical.com/videos/101816-vineland3/Vineland-3-Overview-Webinar-Handout-101816.pdf. [Accessed].

SHROUT PE & FLEISS JL 1979 Intraclass correlations: uses in assessing rater reliability. Psychological bulletin, 86, 420. [PubMed: 18839484]

SLOMINE BS, SILVERSTEIN FS, CHRISTENSEN JR, HOLUBKOV R, TELFORD R, DEAN JM & MOLER FW 2018 Neurobehavioural outcomes in children after In-Hospital cardiac arrest. Resuscitation, 124, 80–89. [PubMed: 29305927]

SPARROW SS, CICCHETTI DV & SAULNIER CA 2016 Vineland Adaptive Behavior Scales, Third Edition, San Antonio, TX, Pearson.

SZATMARI P, GEORGIADES S, DUKU E, BENNETT TA, BRYSON S, FOMBONNE E, MIRENDA P, ROBERTS W, SMITH IM & VAILLANCOURT T 2015 Developmental trajectories of symptom severity and adaptive functioning in an inception cohort of preschool children with autism spectrum disorder. JAMA psychiatry, 72, 276–283. [PubMed: 25629657]

TASSÉ MJ, SCHALOCK RL, BALBONI G, BERSANI JR H, BORTHWICK-DUFFY SA, SPREAT S, THISSEN D, WIDAMAN KF & ZHANG D 2012 The construct of adaptive behavior: Its conceptualization, measurement, and use in the field of intellectual disability. American journal on intellectual and developmental disabilities, 117, 291–303. [PubMed: 22809075]

VAGO C, BULGHERONI S, USILLA A, BIASSONI V, SERRA A, GENTILE S, AJOVALASIT D, LEONARDI M, MASSIMINO M, FIDANI P & RIVA D 2011 Adaptive functioning in children in the first six months after surgery for brain tumours. Disabil Rehabil, 33, 953–60. [PubMed: 21114385]

WASSERSTEIN RL, SCHIRM AL & LAZAR NA 2019 Moving to a World Beyond “p < 0.05”. The American Statistician, 73, 1–19.

WORLD HEALTH ORGANIZATION 2018 International statistical classification of diseases and related health problems (11th Revision).

ZOU G 2012 Sample size formulas for estimating intraclass correlation coefficients with precision and assurance. Statistics in medicine, 31, 3972–3981. [PubMed: 22764084]

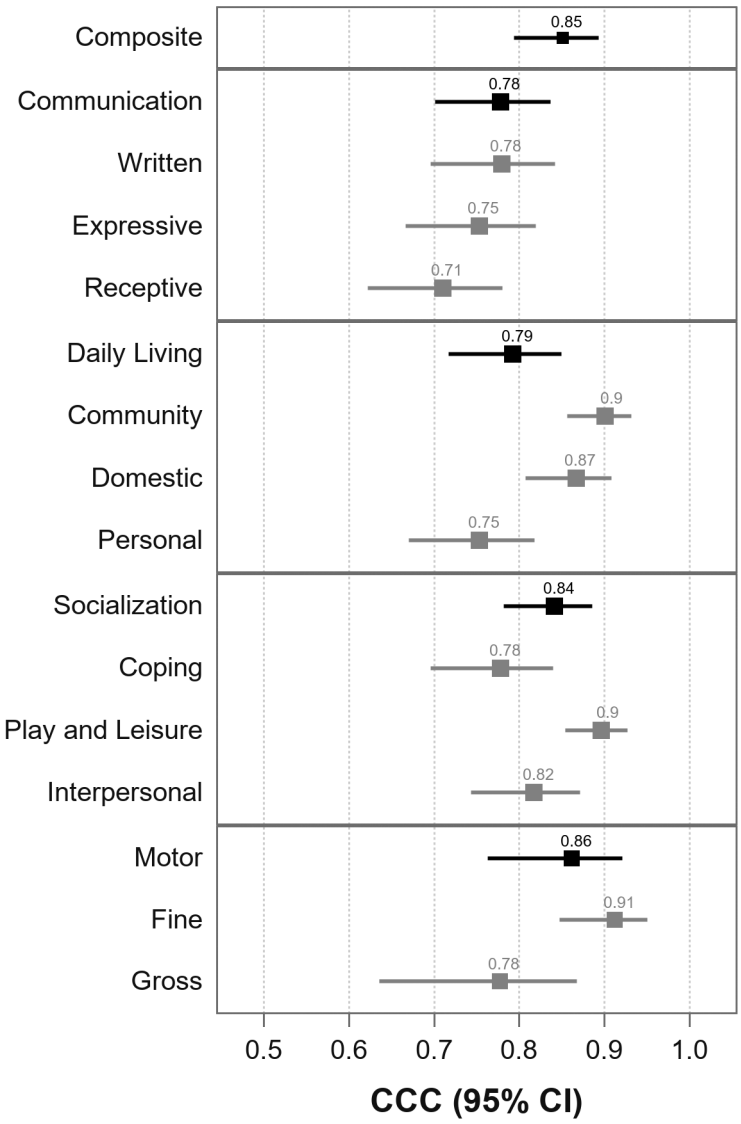

Figure 1.

Agreement between editions, controlling for site and age. CCC = concordance correlation coefficient. CCC for domains (standard scores) are shaded black while subdomains (*v*-scale scores) are shaded gray.

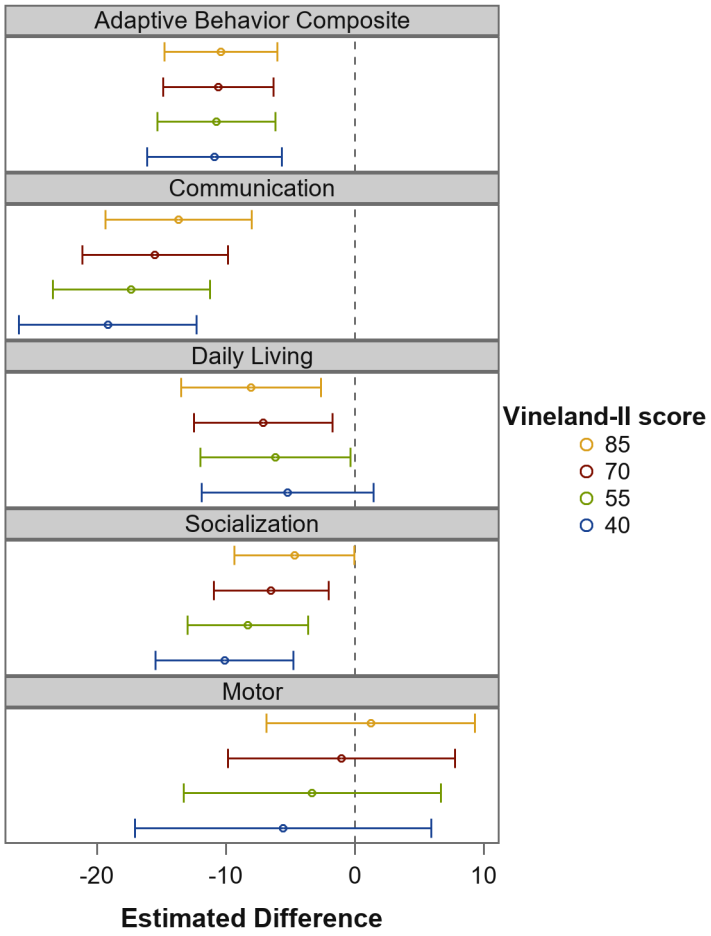

Figure 2.

Estimated mean differences between Vineland-II and Vineland-3 domain scores by level of Vineland-II score, controlling for site and age. figure illustrates the estimated difference between a given Vineland-II score and the predicted Vineland-3 value, controlling for site and age (held constant at 5 years). See also Table 2.

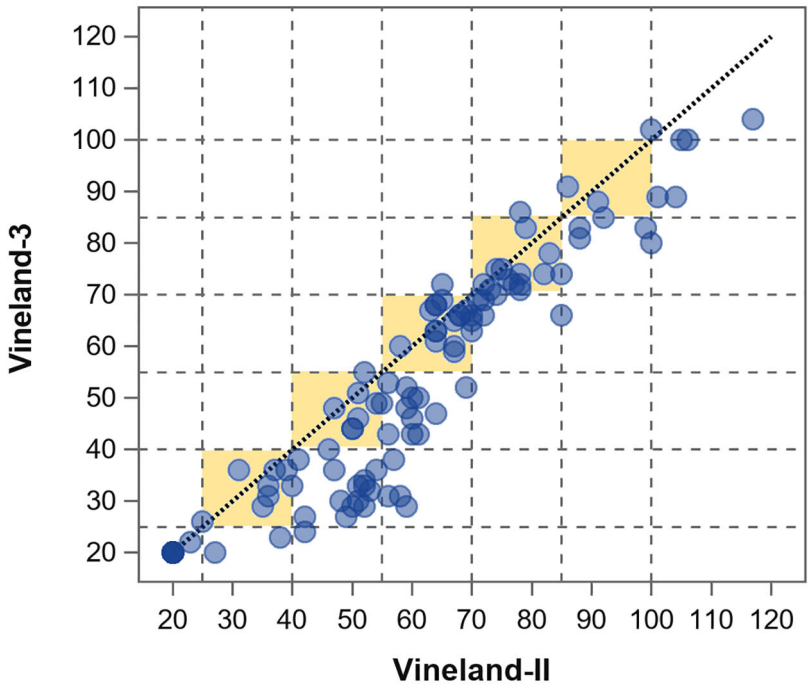

Figure 3.

Raw Adaptive Behavior Composite data. Dotted lines are for reference and demarcate commonly used levels of intellectual disability (<25 = Profound, 25–39 = Severe, 40–54 = Moderate, 55–74 = Mild, 75–84 = Borderline, ≥85 = Average). The diagonal dotted reference line reflects a perfect 1:1 relationship; individuals above this diagonal line had higher Vineland-3 scores, while individuals below the line had higher Vineland-II scores. Shaded areas of graph capture individuals falling within the same group classification (e.g., mild or moderate ID); individuals whose classifications shifted fall outside of the shaded areas. See Table S2 for intercepts and slopes of the Adaptive Behavior Composite and all other subscales.

J Intellect Disabil Res. Author manuscript; available in PMC 2020 January 03.

Table 1.

Participant demographics (N = 106).

	NIH (n = 87)					STAR (n = 19)				
	n	***Mean***	***SD***	***Min***	***Max***	***n***	***Mean***	***SD***	***Min***	***Max***
Male	43 (49%)					14 (74%)				
Age (years)	87 (100%)	12.65	10.47	1.25	50.17	19 (100%)	16.71	13.01	1.67	49.58
Diagnosis[a]										
ASD	2 (2%)					6 (32%)				
ID	63 (72%)					3 (16%)				
ASD + ID	7 (8%)					3 (16%)				
Vineland-II ABC	87 (100%)	61.43	20.50	20	117	18 (95%)	61.33	25.38	20	100
Vineland-3 ABC	87 (100%)	54.70	22.15	20	104	18 (95%)	52.17	24.11	20	83
FS IQ/DQ ≥ 85	10 (11%)					7 (37%)				
FS IQ/DQ 70 – 85	10 (11%)					2 (11%)				
FS IQ/DQ 55 – 69	19 (22%)					0				
FS IQ/DQ 40 – 54	20 (23%)					0				
FS IQ/DQ 25 – 39	5 (6%)					0				
FS IQ/DQ < 25	14 (16%)					2 (11%)				
FS IQ/DQ missing	9 (10%)					6 (32%)				
Time between administrations										
Same day	82 (94%)					16 (84%)				
1–5 days	4 (5%)					3 (16%)				
6–7 days	1 (1%)					0				
Respondent										
Mother	77 (89%)					17 (89%)[b]				
Father	10 (11%)					2 (11%)				
Same clinician administered both forms	60 (69%)					3 (16%)				

[a]Diagnostic checklists were only completed for only autism spectrum disorder (ASD) and intellectual disability/global developmental delay (ID),with global developmental delay diagnosed in children under 5. ID + ASD is mutually exclusive of the singular diagnoses.

[b]For three administrations, the father figure was also present for the interview.

Note: ABC = Adaptive Behavior Composite. FSIQ/DQ = full scale intelligence quotient or developmental quotient. The total number of raters was five at NIH and five at STAR.

Table 2.

Concordance correlation coefficients estimated by variance components analysis, controlling for age and site

Domain	Subscale	CCC	CCC 95% CI lower bound	CCC 95% CI upper bound
Adaptive Behavior Composite	**Adaptive Behavior Composite**	0.85	0.79	0.89
Communication	Receptive	0.71	0.62	0.78
	Expressive	0.75	0.67	0.82
	Written	0.78	0.70	0.84
	Communication	0.78	0.70	0.84
Daily Living Skills	Personal	0.75	0.67	0.82
	Domestic	0.87	0.81	0.91
	Community	0.90	0.86	0.93
	Daily Living Skills	0.79	0.72	0.85
Socialization	Interpersonal Relationships	0.82	0.74	0.87
	Play and Leisure	0.90	0.85	0.93
	Coping Skills	0.78	0.70	0.84
	Socialization	0.84	0.78	0.89
Motor Skills	Gross Motor	0.78	0.64	0.87
	Fine Motor	0.91	0.85	0.95
	Motor Skills	0.86	0.76	0.92

Note: CCC = concordance correlation coefficient; CI = confidence interval