Psychological Reports, 2011, 109, 2, 626-634.

COMPARISON OF THE VINELAND ADAPTIVE BEHAVIOR SCALES, SECOND EDITION, AND THE BAYLEY SCALES OF INFANT AND TODDLER DEVELOPMENT, THIRD EDITION[1]

DOROTHY SCATTONE, DONALD J. RAGGIO, AND WARREN MAY

University of Mississippi Medical Center

Summary.—The Vineland Adaptive Behavior Scales, Second Edition (Vineland–II), and Bayley Scales of Infant and Toddler Development, Third Edition (Bayley–III) were administered to 65 children between the ages of 12 and 42 months referred for developmental delays. Standard scores and age equivalents were compared across instruments. Analyses showed no statistical difference between Vineland–II ABC standard scores and cognitive levels obtained from the Bayley–III. However, Vineland–II Communication and Motor domain standard scores were significantly higher than corresponding scores on the Bayley–III. In addition, age equivalent scores were significantly higher on the Vineland–II for the fine motor subdomain. Implications for early intervention are discussed.

Identifying children with developmental disabilities is the first step toward early intervention services. The Vineland Adaptive Behavior Scales, Second Edition (Vineland–II; Sparrow, Cicchetti, & Balla, 2005), and Bayley Scales of Infant and Toddler Development, Third Edition (Bayley–III; Bayley, 2005), are often considered gold standards in terms of identifying developmental delays in children. A measure of adaptive behavior is required by special education law (*Individuals with Disabilities Education Act Amendments*, 1999; *Individuals with Disabilities Education Improvement Act*, 2004) when assessing intellectual impairment and is strongly recommended when assessing other developmental disabilities as well. Sparrow, *et al.*, defined adaptive behavior as an individual's typical daily performance with respect to communication, daily living, socialization, and motor skills. The Vineland–II is widely used for diagnostic evaluations, monitoring progress, and planning of programs.

The first edition, the Vineland Social Maturity Scale (VSMS), was developed by Doll in 1965, revised by Sparrow, Balla, and Cicchetti (1984), and renamed the Vineland Adaptive Behavior Scales (VABS). These measures were developed out of concern over the use of intelligence tests as a single measure of mental retardation. The latest revision, the Vineland–II (Sparrow, *et al.*, 2005) is a substantial revision over previous editions with expanded age ranges and additional items. It can be administered to the caregiver of the individual being assessed for ages ranging from birth through 90 years. Standard scores ($M = 100$, $SD = 15$) from Communica-

[1]Address correspondence to Dorothy Scattone, University of Mississippi Medical Center, Department of Pediatrics, 2500 N. State Street, Jackson, MS 39216 or e-mail (DScattone@umc.edu).

DOI 10.2466/03.10.PR0.109.5.626-634 ISSN 0033-2941

tion, Daily Living, Socialization, and Motor Skills domains can be combined to create an Adaptive Behavior Composite (ABC). Each domain is further divided into subdomains (see Table 1) that provide age equivalent scores with good test-retest reliability overall and internal consistency on subdomains (Sparrow, *et al.*, 2005).

Whereas the Vineland–II is administered to a parent or caregiver, the Bayley–III is administered by a trained examiner directly to the child being assessed. The Bayley Scales of Infant Development (BSID) was first developed in 1969, revised in 1993 (Bayley Scales of Infant Development-Second Edition, BSID–II), and again in 2005 (Bayley–III). It can be administered to children who are between the ages of 1 and 42 months. Developmental areas assessed by the scales include Cognitive, Language, Motor, Social-Emotional, and Adaptive. Social-Emotional and Adaptive are presented in the form of a questionnaire and are optional. Cognitive, Language, and Motor domains each have a mean of 100 and standard deviation of 15. The Language and Motor domains are further divided into subdomains (see Table 1) that produce age equivalent scores. The Bayley–III has also demonstrated adequate validity and reliability (Bayley, 2005).

Both instruments have advantages and disadvantages. The Vineland–II can be administered without the presence of the individual being assessed and takes into account typical performance, whereas the Bayley–III provides an opportunity for direct interaction with the child. Each instrument, however, has limitations; for example, responses on the Vineland–II may be subject to inaccurate reporting by the caregiver, while responses

TABLE 1
VINELAND–II AND BAYLEY–III DOMAINS AND SUBDOMAINS
COMPARING MEASURES OF EQUIVALENT CONSTRUCTS

Vineland–II	Bayley–III
Communication	Language
Expressive	Expressive communication
Receptive	Receptive communication
Written	
Motor Skills	Motor
Fine	Fine motor
Gross	Gross motor
Daily Living Skills	Cognitive
Personal	
Domestic	
Community	
Socialization	
Interpersonal	
Play and leisure	
Coping	
Adaptive Behavior Composite	

on the Bayley–III may be influenced by rapport building, testing environment, and fatigue on the part of the child being assessed.

Although previous studies have compared cognitive levels with adaptive behavior for children with developmental disorders, many of these studies derived cognitive levels from a variety of instruments (e.g., Liss, Harel, Fein, Allen, Dunn, Feinstein, *et al*., 2001; Bolte & Poustka, 2002; Perry, Flanagan, Geier, & Freeman, 2009). In addition, many of the samples were comprised of individuals approaching adolescence and older (e.g., Liss, *et al*., 2001; Paul, Miles, Cicchetti, Sparrow, Klin, Volkmar, Coflin, & Booker, 2004; Klin, Saulnier, Sparrow, Cicchetti, Volkmar, & Lord, 2007). Only a handful of studies have directly examined adaptive behavior and cognitive skills in very young children, specifically those below age four whose learning and skill development are most at risk (Raggio & Massingale, 1990; Raggio & Massingale, 1993; Raggio, Massingale, & Bass, 1994). Even fewer studies have compared the Vineland Adaptive Behavior Scales with the Bayley Scales of Infant and Toddler Development (Ray-Subramanian, Huai, & Weismer, 2010). The first published study was conducted by Erickson, Johnson, and Campbell (1970) who examined the relationship among scores on the VSMS (Doll, 1965), BSID (Bayley, 1969), and the Cattell Infant Intelligence Scale (Cattell, 1940). The authors found significantly higher adaptive scores on VSMS than cognitive levels from the Cattell and Social Quotients from the BSID.

Raggio and Massingale (1993) compared the VSMS, VABS, and the Mental Scale of the BSID with 44 infants who were referred for developmental delays. Their results suggested a significantly higher ABC on the VABS than either the VSMS Social Quotient or the BSID Mental Development Index. In another study Raggio, *et al.* (1994) compared the VABS-ABC standard scores, VABS age-equivalent scores, and the BSID Mental Development Index. The mean VABS–ABC standard score was significantly higher than age equivalents and the BSID Mental Development Index. The authors suggested these findings may be the result of floor effects with standard scores; therefore, age equivalents may be a better predictor of developmental outcome.

Recently Ray-Subramanian, *et al.* (2010) examined adaptive behavior and cognitive levels for 125 toddlers with autism. The authors found significant delays in adaptive functioning for these 2-year-olds with autism and a group profile consisting of Motor Skills > Daily Living Skills > Socialization > Communication. In addition, Vineland–II domain standard scores and subdomain scores were significantly correlated with the Bayley–III Cognitive domain standard scores. However, the Language and Motor domains of the Bayley–III were not administered. These scales would have provided the opportunity for further comparisons regarding parent report and actual observed skills.

Most of the above studies were conducted with prior versions of the Vineland Adaptive Behavior Scales and the Bayley Scales of Infant and Toddler Development. The only existing study to date comparing the two scales did not include the Language and Motor domains of the Bayley–III. Therefore, the present study was undertaken to (1) replicate Erickson, *et al.* (1970) and Raggio and Massingale (1993) using the latest revision of each instrument and (2) to expand upon Ray-Subramanian, *et al.* (2010) by comparing the Language and Communication and Motor domains of each instrument.

The present investigation addressed the following: (1) the relationship between cognitive level and adaptive behavior for children referred for developmental delays; (a) the relationship between the Vineland–II Communication and Motor domains and Bayley–III Language and Motor domain standard scores; and (b) the relationship between age-equivalent scores on the Vineland–II and Bayley–III for Expressive, Receptive, Gross Motor, and Fine Motor subdomains.

Method

Participants

Sixty-five children between the ages of 12 and 42 months participated in the study. Of these, 47 were boys, and 18 were girls; 30 were African American, and 35 were Euro-American. The mean age was 33 months. Evaluations were conducted by the first author who is a licensed psychologist in a university hospital-based clinic that specializes in diagnosis and treatment of developmental disorders. The children were referred by their primary care physician to the hospital for evaluation of suspected developmental delays.

Instruments

The Vineland–II Survey Interview Form was administered to the primary caregiver of the child being assessed. Immediately following the Vineland–II administration, the Bayley–III was administered directly to the child. The Vineland–II was administered first so that the child's performance on the Bayley–III did not influence caregiver responses on the Vineland–II as the objective was to obtain responses on the child's typical performance at home. Each test was administered and scored by the first author and sessions lasted between 2.5 to 3 hours.

Data Analysis

Correlations and paired *t* tests for the Vineland–II ABC and Bayley–III Cognitive standard scores were performed. In addition, correlations and paired *t* tests were computed between Language and Motor domain standard scores and subdomain age equivalents on both instruments.

The Bland-Altman method (Bland & Altman, 1986) was used to measure agreement on within-subject standard deviation estimates between these two instruments.

Results

Table 2 presents mean standard scores, mean age-equivalent scores, and standard deviations for the Vineland–II and Bayley–III domain and subdomains. The mean Cognitive age-equivalent score on the Bayley–III was 21 months ($SD = 6.7$) in comparison to the mean chronological age of 33 months ($SD = 7.08$). Vineland–II ABC and Communication and Bayley–III Cognitive, Language, and Motor mean standard scores were at least 1.5 standard deviations below the standardization sample mean. The mean ABC standard score on the Vineland–II ranged from 51 to 102 with a mean of 75.3 ($SD = 10.8$). The Vineland–II mean domain standard score was 72.3 ($SD = 14.4$) for Communication and 83.9 ($SD = 12.08$) for Motor. The Bayley–III Cognitive standard score ranged from 55 to 105 with a mean of 76.4 ($SD = 11.4$). The mean domain standard score on the Bayley–III was 67.3 ($SD = 16.6$) for Language and 77.2 ($SD = 16.4$) for Motor.

Correlations were calculated between Vineland–II and Bayley–III domain and subdomain scores and found to be positive and significant at the .05 level (see Table 3). Using the Cicchetti, Koenig, Klin, Volkmar, Paul, and Sparrow (2010) effect size index for correlations (0.10 = trivial; 0.10–0.29 = small; 0.30–0.49 = medium; 0.50–0.69 = large; ≥0.70 = very large), showed a medium correlation for Vineland–II ABC and Bayley–III Cognitive standard scores and a very large correlation for the Vineland–II Communication and Bayley–III Language domain standard scores and age equivalents for receptive, expressive, gross motor, and fine motor.

The relationship between adaptive behavior and cognitive levels was of primary interest. Paired *t* tests yielded no significant difference between the Vineland–II ABC and Cognitive standard score from the Bayley–III. Although differences were not significant with respect to group

TABLE 2
Descriptive Statistics For the Vineland–II and Bayley–III With Comparisons Between Equivalent Domains

Vineland–II	*M*	*SD*	Bayley–III	*M*	*SD*	t_{64}	*p*	Cohen's *d*
ABS, SS	75.3	10.8	Cognitive, SS	76.3	11.3	−0.93		−0.10
Communication, SS	72.3	14.4	Language, SS	67.3	16.6	3.41	.0114	0.03
Motor skills, SS	83.9	12.1	Motor, SS	77.2	16.6	4.10	1.17×10^{-4}	0.47
Receptive, AE	17.3	9.4	Receptive, AE	15.9	19.1	1.52		0.15
Expressive, AE	17.3	8.3	Expressive, AE	16.5	9.6	1.09		0.08
Gross motor, AE	25.7	9.1	Gross motor, AE	24.0	9.5	1.74		0.18
Fine motor, AE	25.2	9.2	Fine motor, AE	21.7	9.4	4.02	1.56×10^{-4}	0.38

Note.—SS = standard score, AE = age equivalent in months.

TABLE 3
VINELAND–II AND BAYLEY–III CORRELATIONS

Vineland–II	Bayley–III	Correlation	95%*CI*
ABC	Cognitive	.64	.47, .76
Communication	Language	.71	.82, .84
Motor	Motor	.61	.74, .78
Receptive	Receptive	.70	.56, .81
Expressive	Expressive	.82	.72, .89
Gross motor	Gross motor	.66	.50, .78
Fine motor	Fine motor	.72	.57, .81

means, there were rather large intra-individual differences. Specifically, the Bland-Altman method of agreement indicated ABC and Cognitive scores varied widely by as much as 20 points above and below the points below the mean for individual cases.

Secondly, the relationship between the two instruments regarding Communication and Motor domains standard scores was examined. The Vineland–II Communication ($t = 3.4$, $p < .001$) and Motor ($t = 4.1$, $p < .001$) domain standard scores were significantly higher than corresponding scores from the Bayley–III for both domains. The Bland-Altman limits for Language and Motor standard scores for the two instruments also varied widely by as much as 28 points above and below the mean for some individuals.

Finally, age equivalents for expressive, receptive, gross motor, and fine motor subdomains were compared for the Vineland–II and Bayley–III. Paired *t* tests showed that Vineland–II age equivalents were significantly higher for fine motor subdomain ($t = 4.02$, $p < .001$) as shown in Table 2, with variability for individual cases ranging from 16 to 18 months above and below the mean.

DISCUSSION

Results suggest that Vineland–II ABC standard scores were statistically similar to Bayley–III cognitive scores, which differs from Erickson, *et al.* (1970) and Raggio and Massingale (1993), who found ABC standard scores higher than cognitive scores. However, in the current study, large within-subject variations were observed—as much as two standard deviations above and below the mean for some individuals. Upon further examination among domains and subdomains, Vineland–II Communication and Motor composite standard scores were found to be significantly higher than Bayley–III Language and Motor composite standard scores, suggesting that more children would be eligible for services using the Bayley–III Language and Motor scales. In addition, age equivalents on the Vineland–II were significantly higher than those from the Bayley–II for the Fine Motor subdomain only. Findings are in agreement with Ray-Sub-

ramanian, *et al.* (2010) who also found Vineland–II and Bayley–III domain and subdomain scores to be highly correlated.

The study is limited by several factors. First, directly testing toddlers was challenging at times and successful administration of the Bayley–III relied heavily on rapport building and adequate energy level on the part of the child being assessed. Taking breaks between subtests on the Bayley–III was necessary for some children as the entire session was somewhat lengthy (i.e., between 2.5 and 3 hours). There were limitations to the Vineland–II administration as well including the possibility of inaccurate reporting on the part of the caregiver. For example, determining whether their child used pronouns, used verbs ending in "ing," or followed one- or two-step commands beyond the specific examples given required lengthy probing for some parents. Other parents began testing the child themselves by asking him or her to "point to your head," for example, suggesting they may not have known the extent of the skills in their child's repertoire. If the Vineland–II is to be administered without the presence of the child, this could be problematic and lead to inaccurate responses for some children.

Data were gathered primarily by the first author through clinical practice rather than a research-funded grant. Thus, the sample may not have been representative; most of the families served in the clinic fall within the lower socioeconomic group. The instruments were not administered in a counterbalanced order. It was necessary to try to obtain parents' perceptions of their children's typical performance. The sample size was relatively small and did not lend itself to comparisons across cognitive levels and differential diagnoses. Floor effects cannot be ruled out. The lowest possible standard score on the Bayley–III is 55. Although only four children in the sample obtained a score of 55, it's possible their real scores could have been lower. Although age equivalents are not usually normed on the standardization sample, they may be a better measure for children whose standard scores are ≤ 55. In addition, age equivalent scores are generally more easily understood by parents.

Future research should also include examining cognitive levels in comparison to all four subdomains on the Vineland–II (Communication, Daily Living, Socialization, Motor) with the new Vineland–II to further explore the relationship among IQ and adaptive behavior, with a larger sample size, information on parental education, socioeconomic status, and comparisons across diagnostic subgroups. In addition, whether or not parents are accurately reporting their children's skills is a factor that should be investigated. Several direct questions on each measure overlap and could be evaluated item by item. For example, both scales evaluate use of pronouns, prepositions, plurals, labeling body parts, labeling

colors, cutting with scissors, holding a crayon properly, walking up and down stairs alternating feet, catching a ball, completing puzzles, running, hopping on one foot, for example. An item-by-item analysis of these specific questions may help to answer this question.

Implications

In conclusion, a measure of cognitive functioning and adaptive behavior are necessary components when diagnosing developmental delays in children. Although the Vineland–II and Bayley–III are often used for this purpose, the results suggest more children would qualify for services using the Bayley–III standard scores for Language and Motor domains. These findings highlight the importance of using multiple measures when making placement decisions and that relying solely on one measure could exclude some children from services.

REFERENCES

Bayley, N. (1969) *Bayley Scales of Infant Development*. San Antonio, TX: The Psychological Corporation.

Bayley, N. (1993) *Bayley Scales of Infant Development.* (2nd ed.) San Antonio, TX: The Psychological Corporation.

Bayley, N. (2005) *Bayley Scales of Infant and Toddler Development.* (3rd ed.) San Antonio, TX: The Psychological Corporation.

Bland, J. M., & Altman, D. G. (1986) Statistical methods for assessment agreement between two methods of clinical measurement. *Lancet*, 327, 307-310.

Bolte, S., & Poustka, F. (2002) The relation between general cognitive level and adaptive behavior domains in individuals with autism with and without co-morbid mental retardation. *Child Psychiatry and Human Development*, 22, 165-172.

Cattell, P. (1940) *The measurement of intelligence of infants and young children*. New York: The Psychological Corporation.

Cicchetti, D. V., Koenig, K., Klin, A., Volkmar, F. R., Paul, R., & Sparrow, S. (2010) From Bayes through marginal utility to effects sizes: a guide to understanding the clinical and statistical significance of the results of autism research findings. *Journal of Autism and Developmental Disorders*. DOI:10,1007/s10803-010-1035-6.

Doll, E. A. (1935, 1965) *Vineland Social Maturity Scale*. Circle Pines, MN: American Guidance Service.

Erickson, M. T., Johnson, N. M., & Campbell, T. A. (1970) Relationship among scores on infant tests for children with developmental problems. *American Journal of Mental Deficiency*, 75, 102-104.

Public Law 105-17 (Individual with Disabilities Education Act Amendments of 1997). (1999) Washington, DC: U.S. Department of Education and the office of Special Education Programs and Rehabilitative Services.

Public Law 108-446 (Individuals with Disabilities Education Improvement Act of 2004). (2004) Washington, DC: U.S. Department of Education and the office of Special Education Programs and Rehabilitative Services.

Klin, A., Saulnier, C. A., Sparrow, S., Cicchetti, D. V., Volkmar, F. R., & Lord, C. (2007) Social and communication abilities and disabilities in higher functioning individuals with autism spectrum disorders: the Vineland and the ADOS. *Journal of Autism and Developmental Disorders,* 37, 748-759.

Liss, M., Harel, B., Fein, D., Allen D., Dunn, M., Feinstein, C., Morris, R., Waterhouse, L., & Rapin, I. (2001) Predictors and correlates of adaptive function in children with developmental disorders. *Journal of Autism and Developmental Disorders,* 31, 219-230.

Paul, R., Miles, S., Cicchetti, D. V., Sparrow, S. S., Klin, A., Volkmar, F. R., Coflin, M., & Booker, S. (2004) Adaptive behavior in autism and pervasive developmental disorder not otherwise specified: microanalysis of scores on the Vineland Adaptive Behavior Scales. *Journal of Autism and Developmental Disorders*, 34, 223-228.

Perry, A., Flanagan, H. E., Geier, J. D., & Freeman, N. L. (2009) Brief report: the Vineland Adaptive Behavior Scales in young children with autism spectrum disorder at different cognitive levels. *Journal of Autism and Developmental Disorders*, 39, 1066-1078.

Raggio, D. J., & Massingale, T. W. (1990) Comparability of the Vineland Social Maturity Scale and the Vineland Adaptive Behavior Scale–Survey Form with infants evaluated for developmental delay. *Perceptual and Motor Skills*, 71, 415-418.

Raggio, D. J., & Massingale, T. W. (1993) Comparison of the Vineland Social Maturity Scale, the Vineland Adaptive Behavior Scales–Survey Form, and the Bayley Scales of Infant Development with infants evaluated for developmental delay. *Perceptual and Motor Skills*, 77, 931-937.

Raggio, D. J., Massingale, T. W., & Bass, J. D. (1994) Comparison of Vineland Adaptive Behavior Scales–Survey Form age equivalent and standard score with the Bayley Mental Development Index. *Perceptual and Motor Skills,* 79, 203-206.

Ray-Subramanian, C. E., Huai, N., & Weismer, S. E. (2010) Brief report: adaptive behavior and cognitive skills for toddlers on the autism spectrum. *Journal of Autism and Developmental Disorders.* DOI:10.1007/s10803-010-1083-y.

Sparrow, S. S., Balla, D. A., & Cicchetti, D. V. (1984) *Vineland Adaptive Behavior Scales*. Circle Pines, MN: American Guidance Service.

Sparrow, S. S., Cicchetti, D. V., & Balla, D. A. (2005*) Vineland Adaptive Behavior Scales.* (2nd ed.) Circle Pines, MN: American Guidance Service.

Accepted September 16, 2011.