"By the Numbers"

**Standardization/Norms**

- **272** Examiners
- **320** children were administered the Social-Emotional Scale
- **1700** children completed the Cognitive, Language and Motor Scale standardization edition
- **44** participating US states and D.C.
- **8** age groups
- **40** in each age group
- **750** children were included in the Adaptive Behavior Scale normative sample during the Vineland™-3 standardization
- **17** age groups
- **100** in each age group
- **9** age groups
- **80-90** in each age group

**Scores Available**

- Scaled Scores
- Standard Scores
- Percentile Ranks
- Confidence Intervals
- Age Equivalents
- Growth Scale Values—with statistical significance
- Discrepancy Analysis—with statistical significance
- Guidance to determine Percent Delay
Reliability

All measures reflect a high degree of reliability in the items, and that the Bayley-4 scales are equally reliable for assessing individuals with different levels of development or individuals with different clinical diagnoses.

<table>
<thead>
<tr>
<th>Cognitive, Language and Motor Sub-tests</th>
<th>Cognitive, Language &amp; Motor Composites</th>
<th>Special Group Studies</th>
<th>Social-Emotional</th>
<th>Adaptive Behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Internal Consistency</strong></td>
<td><strong>.93 to .95</strong></td>
<td><strong>.95 to .96</strong></td>
<td><strong>.97 to .99</strong></td>
<td><strong>.91 to .98</strong></td>
</tr>
<tr>
<td><strong>Test-Retest</strong></td>
<td><strong>.81 to .84</strong></td>
<td><strong>.81 to .85</strong></td>
<td></td>
<td><strong>.72 to .87</strong></td>
</tr>
<tr>
<td><strong>Inter-Rater</strong></td>
<td></td>
<td></td>
<td></td>
<td><strong>.67 to .81</strong></td>
</tr>
</tbody>
</table>

Validity

Correlations/Relationships with Other Measures:

- Bayley-III
- WPPSI®-IV
- PDMS-2
- Information regarding the relation between the Bayley-III and the PLS™-4 and the Bayley–III and the Vineland™-3 is also provided.

Special Group Studies:

- Autism spectrum disorder
- Developmental delay
- Down syndrome
- Language delay and specific language impairment
- Moderate/late premature and very/ extreme premature
- Motor impairment
- Prenatal drug/ alcohol exposure

Classification Accuracy was used in comparing appropriate groups with typically developing peers as measured by the area under the ROC curve. ROC area estimates range from .50 (chance accuracy) to 1.00 (perfect accuracy). Traditional benchmarks suggest that values greater than or equal to .90 are excellent, .80–.89 are good, .70–.79 are fair, and less than .70 are poor.

<table>
<thead>
<tr>
<th>Developmental Delay</th>
<th>Language Delay</th>
<th>Specific Language Impairment</th>
<th>Motor Impairment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classification accuracy of the Cognitive, Language and Motor Scales</td>
<td>.82</td>
<td>.89 (excludes Motor Scale)</td>
<td></td>
</tr>
<tr>
<td>Individual Scales</td>
<td></td>
<td>.86</td>
<td>.80</td>
</tr>
</tbody>
</table>

The Bayley-4 Administration and Technical Manuals include references, detail, and full information on all these data points.