

Bayley-4 Technical Update FAQs



Q. Can you provide additional information about the performance concerns found in the Bayley Scales of Infant and Toddler Development™ (4th ed.; Bayley™-4) normative data?

A. As part of Pearson's postpublication research, 70 cases were identified with performance concerns on the Cognitive, Language, and Motor Scales. Additional analysis indicated that those cases significantly impacted the normative information in the average to extremely high scaled/standard score range.

Q. Why did Pearson replace the cases with performance concerns instead of removing the cases?

A. Removing cases with performance concerns without replacing the cases could have disproportionately impacted certain age groups. Additionally, replacing the 70 cases is a better representation of the true ability distribution of examinees on the Cognitive, Language, and Motor Scales.

Q. Why did Pearson replace the cases with performance concerns with the U.K. validation study cases?

A. There are two primary reasons. First, the Bayley-4 development included several procedures designed to increase the test's utility to assess cognitive, language, and motor development in children regardless of geography. For instance, throughout the development process, Pearson sought feedback and item reviews from an advisory panel composed of nationally and internationally recognized experts in cognitive development, speech-language pathology, occupational and physical therapy, social-emotional development, and early identification of developmental delay in infants and toddlers. Additionally, Pearson sought feedback from test developers and consultants from Pearson's local and international offices on the test content, manipulatives, and scoring criteria. Second, the U.K. version of the Bayley-4 utilizes the same test content, manipulatives, and scoring criteria as the U.S. version with minor adaptations to account for differences in American and British English.

Q. Did Pearson examine the comparability of the items between the U.S. and U.K. versions of the Bayley-4?

A. Yes, using the U.K. validation study sample (n = 337) and a demographically matched sample from the U.S. Bayley-4 standardization edition, both traditional Mantel-Haenszel (Holland & Thayer, 1988) and Rasch bias analyses (Hambleton, 1993) were used to examine the comparability of the U.S. and U.K. versions of the Bayley-4. Items identified as showing differential item functioning underwent additional expert review to identify any potential issues. After completing this review process, the U.K. version of the Bayley-4 utilized the same test content, manipulatives, and scoring criteria as the U.S. version with minor adaptations to account for differences in American and British English.

Q. How did Pearson examine the score comparability of the Cognitive, Language, and Motor Scales between the U.S. and U.K. versions of the Bayley-4?

A. To compare the raw score differences between the U.S. normative cases and the U.K. validation study cases, a matched sample study was conducted where examinees from both samples were matched on age, parent education level, race/ethnicity, and sex. To be included in the analysis, the examinees from the two samples had to match on all four demographic variables resulting in a final matched sample of $n = 293$. Mean raw score differences between the matched samples ranged from 1.9 raw score points (Gross Motor) to .75 (Cognitive). All Cohen's d values were negligible, with the d values ranging from $-.07$ (Fine Motor) to $-.02$ (Cognitive).

Q. How many of the Bayley-4 U.K. validation study cases were included in the U.S. sample?

A. The updated Cognitive, Language, and Motor Scales normative sample included 70 children who participated in the Bayley-4 U.K. validation study.

Q. How were the Bayley-4 U.K. validation study cases selected to be included in the U.S. sample?

A. The Bayley-4 U.K. validation study cases were only added to the updated Cognitive, Language, and Motor Scales normative sample. The Bayley-4 U.K. validation study cases were selected to align with the U.S. demographic variables (i.e., age, parent education level, race/ethnicity, sex) and to match the demographics of the excluded U.S. normative cases as closely as possible.

Q. How were the Bayley-4 U.K. validation study cases distributed in the updated Cognitive, Language, and Motor Scales normative sample?

A. The Bayley-4 U.K. validation study cases appear across the Bayley-4 age range, with the largest number of cases ($n = 12$) in the 17 months 0 days–19 months 30 days age range.

Q. Were any of the Bayley-4 U.K. validation study cases included in the special group's studies or validity studies?

A. No, the Bayley-4 U.K. validation study cases were not included in the special group's studies or validity studies.