NNAT3

Naglieri Nonverbal Ability Test® **Third Edition**

NNAT3 Special Report:

NNAT3 Lower Levels 2018 Updated Norms

The NNAT3 Lower Level norms were updated in 2018 to provide national norms based on an expanded representative norm sample, and to extend the upper end of the norms age range from 11 years 5 months to 11 years 11 months. There were no changes in the test items or forms, the method of administration or scoring, or the scaled scores. The only change was in the age-norm tables used to convert scaled scores to Naglieri Ability Index (NAI) scores.

The updated norms were based on a nationally representative sample of 85,959 students, ages 4 years 0 months to 11 years 11 months. The sample came from two sources. Most students were drawn from those who took NNAT3 in 2016–2017, whose tests were scored by Pearson (either through online administration or central scoring of paper administrations), and whose schools used the test for universal screening at a grade level and were eligible for participation (i.e., did not prohibit their data from being used). Grade 5 students within the target age range were included so that the sample fully represented the population of 11-year-olds, and their Level E scaled scores were converted to equivalent Lower Levels scaled scores. The remainder of the sample consisted of students in the original NNAT3 norm sample who took the test online as part of the 2014–2015 standardization.

Sampling Procedures

A combination of stratified random sampling and statistical weighting was used to achieve a close match of the normative update sample to the national student population demographic characteristics, as reported in the 2015 Census Data from the National Center for Educational Statistics, United States Department of Education. The demographic characteristics controlled were race/ethnicity, socioeconomic status (SES—percentage of students receiving free/reduced lunch), geographic region, urbanicity, and school type (public or private). All demographic variables were controlled at the school level. The weighting procedure involved random deletion or duplication of cases to achieve alignment to the target distributions of demographic variables at each year of age.

Table 1 displays the demographic characteristics of the NNAT3 2018 normative update sample. Characteristics of this sample were very consistent with the 2015 census in terms of geographic region, SES, urbanicity, ethnicity, type of school, and gender.

Table 1. Demographic Characteristics of the NNAT3 2018 Updated Normative Sample

N = 85,959		Percentage of Total U.S. School Enrollment[1]	Percentage of Students in 2018 Norms Update Sample
Ethnicity			
	African American	13.8	15.0
	Asian	4.8	5.5
	Hispanic	21.5	25.6
	White	50.5	49.2
	Other	5.8	4.7
SES			
	Low	20.9	18.8
	Low-Middle	19.6	19.4
	Middle	19.0	20.5
	High-Middle	19.6	20.5
	High	20.9	20.8
Region			
	Northeast	16.3	16.0
	Midwest	21.1	20.4
	South	38.3	38.8
	West	24.3	24.8
Urbanicity			
	Urban	29.9	30.2
	Suburban	43.0	43.1
	Rural	27.1	26.7
School Type			
	Public	88.5	88.5
	Private/Catholic	11.5	11.5
Gender			
	Female	50.0	49.1
	Male	50.0	50.9

[1] National Center for Educational Statistics, United States Department of Education, 2015 Census Data

Norms Development

The NNAT3 2018 updated norms were developed using *inferential norming* (Wilkins, Rolfhus, Weiss & Zhu, 2005; Zhu & Chen, 2011). The first three *moments* (mean, standard deviation, and skewness) of scaled scores were calculated for each year of age. The moments were plotted across age, and various polynomial regressions, ranging from linear to fifth degree polynomial, were fitted to the moment data. Selecting the best-fitting curve for each moment was based on consistency with underlying theoretical expectations and the pattern of growth curves observed in the normative sample. The selected curves were used to derive estimated population moments by age. Then, those estimates were used to generate a theoretical distribution for each normative age group, yielding an age-based percentile for each scaled score. These percentiles were converted to standard scores on a normal (bell curve) distribution with a mean of 100, a standard deviation of 16, and a range of 40 to 160. The progression of standard scores within and across age groups was then examined, and minor irregularities were eliminated by smoothing. The updated 2018 norm tables present the updated NAI equivalents of scaled scores in 1-month intervals for ages 4 years 0 months to 9 years 11 months, and in 3-month intervals from 10 years 0 months to 11 years 11 months. These norm tables are included as Appendix B in the updated NNAT3 Lower Levels Manual and in the updated NNAT3 Hand Scoring Guide.

The other derived scores—percentile ranks, stanines, and normal curve equivalents—have the same relationship to NAI scores as they did in the original NNAT3 norms. Therefore, the table in the Manual and Hand Scoring Guide that shows the conversion of NAI scores to these other scores is unchanged.

Comparison of NNAT3 2015 and 2018 Updated Lower Levels Norms

Table 2 reports the means, standard deviations, and intercorrelations of NAI scores in the normative update sample, using both the 2015 norms and the 2018 updated norms. As expected, the NAIs based on the two sets of norms were almost perfectly correlated because the primary effect of the normative update was to shift NAI scores downward. At each level, mean NAI scores based on the updated norms were about four and a half points lower than those based on the original (2015) norms. Using the updated norms, 10.9% of the students in the 2018 normative update sample scored 120 or above, very close to the 10.6% expected in a normal distribution.

Table 2. Comparison of NAI Scores from the 2015 and 2018 Updated Norms

2015 Norms				2018 Norms Update			
Level	***N***	**Mean**	***SD***	**Mean**	***SD***	**Difference**	***r***
A	9,393	103.0	16.3	98.6	16.4	4.4	.999
B	11,090	106.0	17.2	101.5	16.9	4.5	.999
C	44,015	103.6	15.9	99.1	15.6	4.5	.999
D	13,956	106.2	16.0	101.8	16.0	4.4	.998
All	78,454	104.3	16.2	99.9	16.0	4.4	.999

Note: The average correlation coefficient across all levels was calculated using Fisher's *z* transformation.

Standard Error of Measurement

The standard error of measurement *(SEM)* indicates the precision of a score, that is, the amount by which it would be expected to vary across administrations (assuming no practice or fatigue effects). The change in the norm-sample standard deviation of NAI scores that resulted from the normative update affected the *SEM* of NAI scores. Table 3 presents the *SEM* of NAI scores at each test level, based on the updated norms and the IRT-based reliabilities reported in Chapter 4 of the Lower Levels Manual. Because those reliabilities were the same on both forms at each level and calculated with reference to the standard deviation of ability scores at each level in the overall norm sample, Table 3 provides a single *SEM* that is applicable to both forms at each level.

Table 3. Standard Error of Measurement of 2018 NAI Scores, by Level

Level	*N*	Mean	*SD*	Reliability	NAI *SEM*
A	1,856	90.9	16.2	0.88	5.6
B	1,687	99.5	17.0	0.88	5.9
C	1,922	94.6	14.4	0.84	5.8
D (Grade 3)	1,103	96.1	13.0	0.80	5.8
D (Grade 4)	1,017	95.5	14.5	0.83	6.0

Note: Calculated using IRT-based reliabilities.

Re-Analysis of Validity Studies

In principle, the correlation of NNAT3 with other tests should have been affected little, if at all, by the normative update. The primary effect expected of the updated norms was a shift of all students' NAI scores, which would have little impact on the rank ordering of students' scores and, therefore, little effect on correlations with other variables. However, so that users could compare average NNAT3 (2018) NAI scores with average scores on other tests, the updated norms were used to re-analyze the NNAT3 correlation studies with NNAT2 and OLSAT 8. Results are presented in Tables 4 and 5.

Table 4. Correlations Between NNAT3 (2018 Updated) and NNAT2 NAI Scores

NNAT2 First							NNAT3 First						
		NNAT2		NNAT3					NNAT2		NNAT3		
Level	*N*	Mean	*SD*	Mean	*SD*	*r*	*N*	Mean	*SD*	Mean	*SD*	*r*	Mean *r*[a]
A	58	104.7	13.8	104.6	12.6	.68	56	99.1	11.6	88.1	11.1	.74	**.80**
B	41	106.8	12.3	110.8	13.4	.70	63	98.7	18.5	95.6	16.4	.75	**.77**
C	37	100.7	12.4	103.1	11.3	.76	67	104.0	14.9	101.5	14.2	.73	**.80**
D							46	97.5	15.2	94.4	14.0	.74	**.78**
												Mean of within-level *r*:	**.79**

[a] Mean, using Fisher's z transformation, of within-sequence correlations after adjusting for range restriction on the first administration (Cohen, Cohen, West, & Aiken, 2003).

Table 5. Correlations Between NNAT3 (2018 Updated) and OLSAT 8 Scores

		OLSAT 8 SAI			NNAT3 NAI		Correlation	
NNAT 3 Level	*N*	Scale	Mean	*SD*	Mean	*SD*	*r*	Adj *r*[a]
A	98	Total	112.7	15.0	109.7	23.4	.55	.41
	98	Verbal	109.7	15.2			.25	.17
	98	Nonverbal	113.5	17.0			.66	.51
B	86	Total	100.5	15.3	97.9	16.3	.54	.53
	86	Verbal	101.2	15.2			.43	.42
	86	Nonverbal	99.0	15.2			.55	.54
C	77	Total	103.4	14.0	100.1	13.5	.57	.64
	77	Verbal	107.4	16.1			.50	.56
	77	Nonverbal	99.2	11.9			.56	.63
D	105	Total	92.5	14.0	94.4	15.1	.58	.60
	105	Verbal	91.8	14.9			.47	.49
	105	Nonverbal	93.8	15.1			.57	.59
Mean[b]		Total						.55
		Verbal						.42
		Nonverbal						.57

[a] Correlation adjusted for range restriction on NNAT3 (Cohen, Cohen, West, & Aiken, 2003)

[b] Using Fisher's *z* transformation

NNAT3 correlated slightly higher with NNAT2 when the updated norms were used (average of .79 versus .78). NNAT3 correlations with OLSAT 8 were unaffected, on average. In the NNAT3–NNAT2 study, the difference in mean NAI scores (averaged across administration sequences) was about 5 points at Level A, but near zero at Levels B and C. At Level D, where only one sequence was given, the difference of 3.1 points could reasonably be attributed to the practice effect, suggesting that NNAT3 and NNAT2 NAI scores were very similar at this level as well. In the NNAT3–OLSAT 8 study, where NNAT3 was always administered first, the effect of practice could not be removed statistically. At Levels A through C, NNAT3 scores were about 3 points lower than OLSAT 8 Total SAI scores, a difference that was consistent with a practice effect and with the Flynn effect. At Level D, NNAT3 scores were slightly higher than OLSAT 8 scores.

References

Wilkins, C., Rolfhus, E., Weiss, L., & Zhu, J. (2005). *A simulation study comparing continuous and traditional norming with small sample sizes.* Paper presented at the annual meeting of the American Educational Research Association, Montreal, Canada.

Zhu, J. & Chen, H. (2011). Utility of inferential norming with smaller sample sizes. *Journal of Psychoeducational Assessment, 29,* 570–580.