

# NNAT<sup>®</sup> 2 Scores with and Without a Time Limit

Pearson Psychometric & Research Services

January 2013

**Pearson, Naglieri Nonverbal Ability Test, and NNAT** are trademarks, in the U.S. and/or other countries, of Pearson Education, Inc. or its affiliate(s).

The *Naglieri Nonverbal Ability Test*<sup>®</sup>—Second edition is designed to evaluate general ability within a time frame of 30 minutes for each of its seven, grade-based levels. Each level contains 48 items, and it has been found that 30 minutes allows most students to finish all items. Some students, however, have IEPs or 504 plans that specify additional time for taking tests such as NNAT2. It is important, therefore, to investigate whether providing additional time would differentially benefit students who are allowed more time to complete the 48 items.

This document is a report of an NNAT2 research study designed to compare the test scores obtained using the current 30-minute time limit and an extended time limit that was not to exceed 60 minutes. Testing took place October 1 through November 2, 2012, at five schools, using the paper version of the NNAT2 levels for Kindergarten, grade 1, and grade 2.

For this study, student scores, raw scores for within grade/age comparison, and z-scores for across-grade comparison were used to determine the effects of additional time on student performance.

Two research questions were addressed in this study. The first research question was whether providing additional time for test-taking increased student performance. Student raw scores were used to indicate the performance on the test. The second research question was whether providing additional time for test-taking would increase the completion rate. Students were said to have completed the test if they attempted the test and responded to the last item. Teachers were asked to record the time each student took to finish the test. The time record was used to evaluate this research question.

For the timing study, the design was implemented as described below:

- The target was 80 students each in three grades: Kindergarten and grades 1 and 2.
- At each grade, students were divided into two groups of 40 students each: group 1 and group 2.
- Each of those groups was further divided into sub-groups of 20 students each: group 1a, group 1b, group 2a, and group 2b.

Two versions, or forms, of NNAT2 were used in the study. Form 1 was the current paper version of NNAT2, and Form 2 was a new version developed in 2012 for a special program. The new form was identical to NNAT2 Form 1 in blueprint and item/form difficulty. The students in each group 1a took Form 2 first with a 30-minute time limit enforced. They then were tested with Form 1 but given extended time for test completion with a maximum limit of 60 minutes enforced. Similarly, students in group 1b took Form 2 first but were given an extended 60-minute period of time in which to complete the test. They then in turn tested with Form 1 with the 30-minute time limit enforced. There was a one-week interval between the two tests for the students.

The same methodology was followed for groups 2a and 2b, except that the order of form administration was reversed so that students tested with Form 1 first and then Form 2. The data collection design and the actual n counts in the timing study are provided in Table 1 below. Note that numbers in the parentheses in the “Time” columns are completion rates, and numbers in the parentheses in the “Mean Raw Score” columns are standard deviations.

Table 1. Data Collection Design, Means, Standard Deviations, and Ns by Group (N = 251).

	Form 1 Time	Form 2 Time	Form 1 Mean Raw Score	Form 2 Mean Raw Score	n
<b>K, Group 1a</b>	F2_30 (100%)	F1_60 (92.6%)	25.74 (6.65)	27.30 (8.12)	27
<b>K, Group 1b</b>	F2_60 (96.3%)	F1_30 (100%)	27.56 (8.15)	28.89 (8.68)	27
<b>K, Group 2a</b>	F1_30 (100%)	F2_60 (82.4%)	18.12 (10.22)	21.47 (11.69)	17
<b>K, Group 2b</b>	F1_60 (93.8%)	F2_30 (93.8%)	26.88 (9.60)	29.75 (7.08)	16
<b>Grade 1, Group 1a</b>	F2_30 (100%)	F1_60 (90.0%)	30.10 (7.19)	34.70 (6.27)	20
<b>Grade 1, Group 1b</b>	F2_60 (94.7%)	F1_30 (94.7%)	29.53 (7.76)	33.05 (6.59)	19
<b>Grade 1, Group 2a</b>	F1_30 (100%)	F2_60 (100%)	28.25 (7.96)	28.50 (9.67)	16
<b>Grade 1, Group 2b</b>	F1_60 (100%)	F2_30 (100%)	31.67 (5.62)	33.10 (5.19)	21
<b>Grade 2, Group 1a</b>	F2_30 (100%)	F1_60 (100%)	29.50 (6.87)	31.50 (6.45)	20
<b>Grade 2, Group 1b</b>	F2_60 (94.4%)	F1_30 (100%)	29.22 (6.32)	30.28 (6.62)	18
<b>Grade 2, Group 2a</b>	F1_30 (100%)	F2_60 (100%)	28.21 (5.65)	31.21 (5.96)	24
<b>Grade 2, Group 2b</b>	F1_60 (100%)	F2_30 (100%)	30.88 (5.76)	31.19 (6.09)	26

Note: F1\_30: Form 1 with 30 minutes, F1\_60: Form 1 with maximum of 60 minutes; F2\_30: Form 2 with 30 minutes, F2\_60: Form 2 with maximum of 60 minutes

From Table 1, a few observations can be made:

1. In general, student raw scores were higher on the second test administered, regardless of time variation and test sequence.
2. The completion rate was very similar between the two timing conditions for students in grades 1 and 2.
3. One of the conditions for grade K students (when Form 2 was given as the second form with 60 minutes time allowed) appeared to have a lower than usual completion rate (82.4%).

**Statistical Analysis**

The counter-balanced design made it possible to use students as their own control when analyzing the test results. In addition, it is reasonable to assume that the effect due to practice (or test sequence) is canceled out due to counter-balancing. As such, a two-way factorial design (form and timing) with emphasis on testing the effect due to time allowed (30 versus 60 minutes) was used by level and with a nested two-way factorial design across levels. The test scores and the completion rate were compared and tested for the effect of time allowed on these two variables. All comparisons were carried out both by grade and across grades. The main effect was tested, and then a one-tailed significance test was performed on the student raw scores and the completion rates. For these analyses, students within each of the 12 conditions were considered to be nested within each condition. Finally, although the form effect was not of specific interest for this study, test form was treated as another factor so that the effect of timing could be gauged more precisely.

**Research Question 1:** Does providing additional time for test-taking increase student performance?

The hypotheses are as follows:

*H<sub>0</sub>*: There is no difference in student raw scores between the 30-minute group and the extended-time group.

*H<sub>1</sub>*: The student raw scores of the extended-time group are greater than those in the 30-minute group.

Comparisons were made first within each grade level and then across all grades. For cross-grade comparison, z-scores were computed within each grade, and the resulting z-scores were used for comparison. The results are provided in Table 2. The findings indicate that the increase of time limit did not result in significant differences between student raw scores.

**Table 2. Comparison of Mean Raw Scores and Standard Deviations by Time.**

Grade	NNAT2 Level	N	Mean Raw Score		F- value	P- value	Effect Size <sup>1</sup>
			30 minutes	60 minutes			
K	A	87	25.97 (9.05)	26.16 (9.33)	0.01	n.s. <sup>2</sup>	0.02
1	B	76	31.28 (7.28)	31.26 (7.56)	0.01	n.s. <sup>2</sup>	<0.01
2	C	88	29.81 (6.27)	30.77 (6.05)	0.63	n.s. <sup>2</sup>	0.16
All		251	-0.03 <sup>3</sup> (0.99)	0.03 <sup>3</sup> (1.00)	0.24	n.s. <sup>2</sup>	0.06

<sup>1</sup>Effect size was computed by first dividing the mean differences by the pooled standard deviation, then taking the absolute value of the result.

<sup>2</sup>n.s. = Not Statistically Significant.

<sup>3</sup>For this comparison, student raw scores (jointly for both timing conditions) were converted to a z-score within each grade. The resulting z-scores from all three grades were then used as the dependent variable.

Research Question #2: Does providing additional time for test-taking increase student completion rate?

$H_1$ : The completion rate of the extended-time group is greater than the 30-minute group.

The hypotheses are as follows:

$H_0$ : There is no difference in completion rate between the 30-minute group and the extended-time group.

Similar to the analysis above, the comparison of completion rates was performed both within each grade and across grades. However, since whether a student responded to the last item was a Yes/No question, it was unnecessary to perform a score transformation.

**Table 3. Completion Rate Comparison by Time.**

Grade	NNAT2 Level	N	Completion rate		F-value <sup>2</sup>	P-value	Effect Size <sup>1</sup>
			30 minutes	60 minutes			
K	A	87	0.99 (0.11)	0.92 (0.27)	4.85	<.05	0.34
1	B	76	0.99 (0.11)	0.96 (0.20)	0.90	n.s. <sup>3</sup>	0.19
2	C	88	1 (0)	0.99 (0.11)	1.11	n.s. <sup>3</sup>	0.13
All		251	0.99 (0.09)	0.96 (0.21)	6.55	<.05	0.19

<sup>1</sup>Effect size was computed by first dividing the mean differences by the pooled standard deviation, then taking the absolute value of the result.

<sup>2</sup>Italicized entries indicates a statistical significance at the 0.05 level.

<sup>3</sup>n.s. = Not Statistically Significant.

As indicated in Table 3 above, the difference in completion rate was mostly attributable to the Kindergarten students, as grades 1 and 2 did not have significantly different completion rates. Because an overall effect of timing condition on the completion rate was found, further analysis was conducted.

After a review of the raw scores and distributions, a follow-up study was performed by excluding the “K group 2a” condition, as that group appeared to be an outlier. Once that group was excluded, the following results were observed:

**Table 4. Completion Rate Comparison by Time with K Group 2a Excluded.**

Grade	NNAT2 Level	N	Completion rate		F-value	P-value	Effect Size <sup>1</sup>
			30 minutes	60 minutes			
All		234	0.99 (0.09)	0.97 (0.18)	3.52	n.s. <sup>2</sup>	0.15

<sup>1</sup>Effect size was computed by first dividing the mean differences by the pooled standard deviation, then taking the absolute value of the result.

<sup>2</sup>n.s. = Not Statistically Significant.

After exclusion of this group, the results indicated that the effect of timing on completion rate was not statistically significant.

higher completion rate compared to their counterparts in the standard 30-minute condition. In combination, these results suggest ultimately that providing additional time does not alter the scores the NNAT2 yields. It is, therefore, acceptable for administrators to provide extra time for students with accommodations as directed by their IEP/504 plan.

### Results and Discussion

Two research questions were addressed in this study. When a comparison of raw scores was made, students taking the paper version of NNAT2 with the 60-minute condition did not score significantly differently from those with the standard 30-minute condition. Further, the students in the 60-minute condition did not have a