

The National Spanish Examinations Reliability

NSE recognizes that test reliability is defined as the degree to which the test gives consistent results each time it is given.

In other words, reliability answers the following questions:

1. Can I depend on the test to measure the same outcomes consistently?
2. Given all the other variables being the same, will the test produce the same results again?

NSE uses the Kuder-Richardson 21 formula to calculate reliability coefficients:

$$r(\text{reliability}) = \frac{(K)(SD^2) - M(K-M)^2}{(SD^2)(K-1)}$$

K = the number of items in the list

SD = the standard deviation of the scores

M = the mean of the scores

The scores of reliability are judged against a perfect score of 1.00. The closer the reliability coefficient is to 1.00, the better it is. Most standardized tests usually have a reliability coefficient of .90 or above.¹

Reliability Coefficients for the 2013 National Spanish Exam

Level 01	Vocabulary	Grammar	Achievement	Reading	Listening	Proficiency	Total
Regular (N=5501)							
Mean Raw Score	69.2	53.1	122.3	62.4	55.4	117.7	240.0
Standard Deviation	18.6	18.0	33.7	20.5	16.7	33.6	62.3
Reliability coefficient	0.948	0.933	0.963	0.954	0.920	0.962	0.978
Outside Experience (N=1183)							
Mean Raw Score	78.9	60.1	139.1	71.3	63.0	134.3	273.4
Standard Deviation	16.5	18.6	32.2	20.0	17.7	34.4	61.4
Reliability coefficient	0.948	0.940	0.964	0.959	0.935	0.968	0.980
Bilingual (N=367)							
Mean Raw Score	88.6	64.8	153.5	75.1	74.6	149.8	303.2
Standard Deviation	15.9	18.6	31.5	22.2	19.9	39.6	64.6
Reliability coefficient	0.970	0.944	0.969	0.972	0.962	0.981	0.985

Level 1	Vocabulary	Grammar	Achievement	Reading	Listening	Proficiency	Total
Regular (N=36205)							
Mean Raw Score	74.1	58.3	132.4	67.4	57.8	125.2	257.6
Standard Deviation	18.7	19.1	35.2	21.5	18.1	36.3	66.1
Reliability coefficient	0.955	0.942	0.969	0.962	0.935	0.969	0.981
Outside Experience (N=3643)							
Mean Raw Score	83.9	66.8	150.7	76.9	68.5	145.4	296.1
Standard Deviation	17.1	18.8	33.5	19.9	19.4	36.4	65.6
Reliability coefficient	0.964	0.947	0.972	0.965	0.952	0.975	0.985
Bilingual (N=1969)							
Mean Raw Score	91.0	68.6	159.6	77.9	74.7	152.7	312.2
Standard Deviation	14.6	17.4	29.3	22.5	22.7	42.6	64.8
Reliability coefficient	0.971	0.938	0.967	0.976	0.973	0.985	0.986

Level 2	Vocabulary	Grammar	Achievement	Reading	Listening	Proficiency	Total
Regular (N=35272)							
Mean Raw Score	51.7	38.7	90.3	64.6	55.7	120.3	210.6
Standard Deviation	18.2	19.3	34.4	19.4	19.5	35.2	63.1
Reliability coefficient	0.934	0.946	0.963	0.949	0.945	0.966	0.977
Outside Experience (N=2879)							
Mean Raw Score	65.6	47.1	112.7	72.2	68.8	141.0	253.8
Standard Deviation	19.6	19.5	35.6	17.3	20.5	34.3	63.9
Reliability coefficient	0.951	0.944	0.966	0.943	0.959	0.969	0.980
Bilingual (N=2712)							
Mean Raw Score	81.1	54.9	136.0	72.5	76.2	148.7	284.7
Standard Deviation	19.6	19.5	36.2	19.7	23.0	39.4	67.4
Reliability coefficient	0.970	0.944	0.972	0.958	0.975	0.980	0.984

Level 3	Vocabulary	Grammar	Achievement	Reading	Listening	Proficiency	Total
Regular (N=28750)							
Mean Raw Score	43.1	46.5	89.6	59.5	52.3	111.8	201.4
Standard Deviation	15.3	19.1	30.4	19.1	20.7	35.8	59.7
Reliability coefficient	0.904	0.941	0.951	0.943	0.951	0.966	0.974
Outside Experience (N=2410)							
Mean Raw Score	52.4	49.8	102.3	63.6	60.3	123.8	226.1
Standard Deviation	19.8	20.0	35.5	19.5	22.8	38.3	67.2
Reliability coefficient	0.946	0.947	0.965	0.949	0.963	0.973	0.981
Bilingual (N=2700)							
Mean Raw Score	77.5	63.0	140.5	66.9	70.7	137.5	278.1
Standard Deviation	21.2	18.0	36.4	20.3	23.0	39.9	66.8
Reliability coefficient	0.971	0.937	0.973	0.956	0.971	0.978	0.983

Level 4	Vocabulary	Grammar	Achievement	Reading	Listening	Proficiency	Total
Regular (N=16776)							
Mean Raw Score	44.6	36.8	81.4	61.5	49.0	110.5	191.9
Standard Deviation	16.4	18.2	30.1	20.9	20.2	36.7	59.6
Reliability coefficient	0.917	0.939	0.951	0.955	0.948	0.968	0.974
Outside Experience (N=1440)							
Mean Raw Score	52.5	38.3	90.8	64.4	56.4	120.8	211.6
Standard Deviation	19.0	19.2	33.8	21.3	21.6	38.8	64.8
Reliability coefficient	0.941	0.946	0.961	0.959	0.957	0.973	0.979
Bilingual (N=2191)							
Mean Raw Score	73.4	49.1	122.6	64.4	60.3	124.7	247.3
Standard Deviation	17.9	18.1	32.7	21.0	22.4	39.1	59.2
Reliability coefficient	0.949	0.933	0.960	0.958	0.962	0.974	0.975

Level 5	Vocabulary	Grammar	Achievement	Reading	Listening	Proficiency	Total
Regular (N=5636)							
Mean Raw Score	43.6	41.4	85.0	66.3	48.9	115.3	200.2
Standard Deviation	15.9	19.1	30.7	20.4	20.6	37.2	60.7
Reliability coefficient	0.912	0.943	0.953	0.956	0.951	0.970	0.975
Outside Experience (N=792)							
Mean Raw Score	51.8	45.3	97.0	67.4	56.3	123.7	220.8
Standard Deviation	19.3	20.3	35.6	20.7	21.9	38.7	66.3
Reliability coefficient	0.942	0.949	0.965	0.959	0.958	0.973	0.980
Bilingual (N=1216)							
Mean Raw Score	70.4	55.3	125.7	65.5	56.5	121.9	247.7
Standard Deviation	18.9	19.2	34.6	20.2	21.3	37.5	60.5
Reliability coefficient	0.951	0.942	0.966	0.954	0.955	0.971	0.977

Level 6	Vocabulary	Grammar	Achievement	Reading	Listening	Proficiency	Total
Regular (N=507)							
Mean Raw Score	49.0	43.1	92.1	69.8	54.0	123.8	215.9
Standard Deviation	18.3	20.9	35.3	21.4	22.1	40.3	67.9
Reliability coefficient	0.935	0.954	0.965	0.964	0.959	0.976	0.981
Outside Experience (N=194)							
Mean Raw Score	55.1	45.2	100.3	70.2	59.7	130.0	230.2
Standard Deviation	18.5	20.9	36.2	18.5	21.7	36.6	64.6
Reliability coefficient	0.937	0.953	0.967	0.949	0.958	0.971	0.979
Bilingual (N=383)							
Mean Raw Score	76.9	61.1	138.0	68.5	59.9	128.5	266.5
Standard Deviation	16.3	17.8	30.9	21.7	22.1	40.5	59.2
Reliability coefficient	0.943	0.935	0.960	0.964	0.960	0.977	0.977

¹John A. Kaufhold, Basic Statistics for Educational Research (New York: iUniverse, Inc., 2007) pp. 43-46.