

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:

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

K = the number of items in the test

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 2010 National Spanish Examinations

Level 01	Vocab	Gram	Ach	Read	List	Prof	Total
Regular (N=4072)							
Mean Raw Score	56.5	50.4	106.9	54.2	52.4	106.6	213.6
Standard Deviation	19.3	18.0	34.6	17.8	18.0	32.3	61.1
Reliability Coefficient	.943	.932	.963	.931	.932	.957	.976
Outside Experience (N=955)							
Mean Raw Score	61.4	59.0	120.4	58.5	59.0	117.4	237.9
Standard Deviation	19.3	19.7	36.2	19.8	20.4	37.2	67.9
Reliability Coefficient	.946	.947	.968	.947	.951	.970	.982
Bilingual (N=232)							
Mean Raw Score	69.1	76.8	145.9	67.9	70.4	138.3	284.2
Standard Deviation	19.9	22.8	40.1	22.6	22.7	42.9	76.0
Reliability Coefficient	.956	.976	.980	.967	.969	.982	.988

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

Level 1	Vocab	Gram	Ach	Read	List	Prof	Total
Regular (N=29169)							
Mean Raw Score	62.9	56.0	118.9	60.0	56.8	116.7	235.7
Standard Deviation	20.2	19.3	36.9	19.0	18.8	34.6	65.8
Reliability Coefficient	.952	.943	.969	.943	.940	.964	.980
Outside Experience (N=2740)							
Mean Raw Score	68.5	67.4	135.9	67.4	66.1	133.5	269.3
Standard Deviation	19.7	20.3	37.2	20.0	20.4	37.8	69.6
Reliability Coefficient	.954	.956	.973	.955	.956	.974	.984
Bilingual (N=1453)							
Mean Raw Score	71.3	81.8	153.1	72.6	72.6	145.2	298.3
Standard Deviation	18.9	20.3	36.0	21.9	21.9	41.2	69.9
Reliability Coefficient	.952	.973	.977	.968	.968	.981	.987

Level 2	Vocab	Gram	Ach	Read	List	Prof	Total
Regular (N=28058)							
Mean Raw Score	37.5	53.1	90.6	63.9	59.6	123.5	214.1
Standard Deviation	17.5	18.9	33.2	19.5	19.5	35.7	62.8
Reliability Coefficient	.933	.940	.960	.949	.946	.968	.977
Outside Experience (N=2363)							
Mean Raw Score	44.3	65.3	109.6	70.7	68.8	139.4	249.0
Standard Deviation	20.0	20.8	37.7	19.7	20.3	37.1	68.5
Reliability Coefficient	.948	.957	.970	.956	.957	.974	.982
Bilingual (N=1859)							
Mean Raw Score	54.5	83.6	138.1	76.9	75.3	152.2	290.3
Standard Deviation	18.4	19.1	34.5	19.6	20.0	36.9	64.1
Reliability Coefficient	.936	.972	.969	.964	.963	.978	.983

Level 3	Vocab	Gram	Ach	Read	List	Prof	Total
Regular (N=24853)							
Mean Raw Score	45.7	54.3	100.0	69.2	58.3	127.5	227.5
Standard Deviation	18.6	18.0	33.4	19.7	19.5	35.5	62.0
Reliability Coefficient	.938	.933	.960	.955	.945	.968	.977
Outside Experience (N=2148)							
Mean Raw Score	48.2	63.2	111.4	72.0	64.3	136.3	247.6
Standard Deviation	20.1	22.0	38.8	19.8	21.1	37.4	70.0
Reliability Coefficient	.948	.961	.972	.958	.958	.974	.983
Bilingual (N=1943)							
Mean Raw Score	59.7	84.9	144.6	73.5	72.0	145.4	290.1
Standard Deviation	17.5	17.8	32.6	20.3	20.9	37.7	60.4
Reliability Coefficient	.931	.969	.967	.962	.963	.977	.981

Level 4	Vocab	Gram	Ach	Read	List	Prof	Total
Regular (N=14759)							
Mean Raw Score	38.3	53.7	92.0	60.7	55.2	115.9	207.9
Standard Deviation	19.9	16.3	32.0	20.8	19.9	36.7	61.6
Reliability Coefficient	.950	.916	.956	.955	.947	.969	.976
Outside Experience (N=1299)							
Mean Raw Score	41.8	60.8	102.6	63.5	58.8	122.3	224.9
Standard Deviation	21.0	18.6	35.8	22.0	21.2	39.4	67.9
Reliability Coefficient	.955	.941	.966	.962	.956	.974	.981
Bilingual (N=1946)							
Mean Raw Score	52.4	78.7	131.1	61.5	59.0	120.5	251.6
Standard Deviation	17.9	15.7	30.4	23.2	21.4	40.3	58.4
Reliability Coefficient	.932	.941	.956	.966	.957	.975	.975

Level 5	Vocab	Gram	Ach	Read	List	Prof	Total
Regular (N=5510)							
Mean Raw Score	52.9	46.8	99.7	62.8	60.2	123.0	222.6
Standard Deviation	20.3	18.1	34.3	20.0	19.9	36.1	62.0
Reliability Coefficient	.949	.933	.962	.951	.949	.968	.977
Outside Experience (N=623)							
Mean Raw Score	53.3	57.0	110.3	64.1	63.9	128.0	238.3
Standard Deviation	20.1	20.2	35.7	19.5	19.0	34.8	59.6
Reliability Coefficient	.948	.950	.966	.949	.945	.967	.975
Bilingual (N=1148)							
Mean Raw Score	71.8	81.9	153.7	61.4	58.9	120.3	274.0
Standard Deviation	16.5	16.3	30.0	21.8	21.5	39.3	55.5
Reliability Coefficient	.935	.954	.965	.960	.957	.974	.974

Level 6	Vocab	Gram	Ach	Read	List	Prof	Total
Regular (N=398)							
Mean Raw Score	58.2	54.6	112.7	65.1	62.8	127.9	240.6
Standard Deviation	21.2	18.6	35.9	20.9	21.5	38.5	64.7
Reliability Coefficient	.956	.937	.967	.958	.959	.974	.980
Outside Experience (N=87)							
Mean Raw Score	52.8	60.1	112.8	62.2	64.7	126.9	239.7
Standard Deviation	20.3	21.4	36.6	24.4	21.6	42.9	70.7
Reliability Coefficient	.949	.957	.968	.970	.960	.980	.983
Bilingual (N=300)							
Mean Raw Score	72.5	84.5	157.0	59.8	55.7	115.6	272.6
Standard Deviation	18.7	16.4	33.2	23.8	24.0	43.8	65.1
Reliability Coefficient	.953	.961	.974	.967	.967	.979	.982