



4-1980

Normal Curve Equivalents in Bilingual Program Evaluation

Samuel Houston

Jon Kayne

Lee Shannon

Follow this and additional works at: <https://commons.und.edu/tl-nirp-journal>



Part of the [Scholarship of Teaching and Learning Commons](#)

Recommended Citation

Houston, Samuel; Kayne, Jon; and Shannon, Lee (1980) "Normal Curve Equivalents in Bilingual Program Evaluation," *Teaching and Learning: The Journal of Natural Inquiry & Reflective Practice*: Vol. 5 : Iss. 1 , Article 5.

Available at: <https://commons.und.edu/tl-nirp-journal/vol5/iss1/5>

This Article is brought to you for free and open access by UND Scholarly Commons. It has been accepted for inclusion in *Teaching and Learning: The Journal of Natural Inquiry & Reflective Practice* by an authorized editor of UND Scholarly Commons. For more information, please contact und.common@library.und.edu.

Normal Curve Equivalent in Bilingual Program Evaluation

Samuel Houston

University of Northern Colorado

Jon Kayne

Hillsdale College

Lee Shannon

Regis College

Introduction

The formative evaluation of bilingual programs poses special problems for the outside evaluator. What are the criterion variables to be analyzed--project objectives, measures in the cognitive domain, measures in the affective area? Are the subjects assigned on a random or matched basis to the experimental and control groups? What about the use of grade--equivalent scores, percentile ranks and T-scores in the data analysis?

Why Normal Curve Equivalent (NCEs)?

One of the most popular and widely used measures of achievement are grade-equivalent scores. Not only do they tend to be grossly inadequate because of the manner in which they are constructed, they are scaled in such a way that the usual methods of statistical averaging are not legitimate. Percentile ranks are not unlike grade-equivalent scores in that statistical averaging should also be avoided because they too do not form an equal-interval scale. While T-scores (standard scores with a mean of 50 and a standard deviation of 10) possess the required features for the usual statistical operations, it is felt they lack meaningfulness for the usual consumers of educational research.

Normal Curve Equivalent (NCE) scores are normalized scores with a mean of 50 and a standard deviation of 21.06; these values were selected because they produce an exact match between NCEs of one and 99 percentile ranks of one and 99. Thus, percentile ranks and NCEs have the same range and the same midpoint but NCEs have an equal-interval scale. A gain of ten NCEs corresponds exactly to the same amount of student performance achievement at the extreme low end of the achievement distribution as it does for performers in the middle range. An important feature of the use of NCEs is that the experimental group serves as its own control group. Its later success in NCE units is measured against the baseline data gathered initially which also are expressed in NCE units.

In the evaluation of a research project in which NCEs are utilized, all NCE gains greater than zero are considered useful. Whenever a particular program shows that an NCE is greater than zero, it is assumed that the participating students profited from the project. While it is not possible to designate any specific NCE gain as a minimal standard for project success, nevertheless a 5-NCE gain produced in an experimental group of 100 pupils might be considered as good as a 10-NCE gain produced in a group of 50 pupils, assuming the same dollars were spent (Tallmadge and Wood, 1976) in both studies.

Grade Two Bilingual Example

An example of the NCE approach is presented for Grade 2 Bilingual Project students selected from Weld County School District RE-5J in Colorado. The data are selected from a 1979 evaluation report of the Bilingual Project at the primary level which also included results from the Kindergarten, Grade 1 and Grade 3.

The Bilingual Project objectives and the extent to which they were satisfied are presented in Table 1. The six Bilingual Project objectives included

TABLE 1
GRADE TWO BILINGUAL DATA (N=44)

		TEST 1 Oral Language	TEST 2 Identi- fication	TEST 3 Part I Reading	TEST 4 Part II Reading	TEST 5 PART III Reading	TEST 6 Total Points
Highest Possible Points	Pre	80	20	25	10	10	145
	Post	80	20	25	10	10	145
Mean (\bar{X}) Standard Deviation (SD)	Pre	10.2	16.1	19.2	5.3	5.9	55.4
	Post	9.2	3.2	4.3	3.1	3.1	16.5
Mean (\bar{X}) Standard Deviation (SD)	Pre	36.7	18.3	22.2	8.2	8.5	95.3
	Post	25.5	2.3	3.5	2.2	1.8	32.0
t-test on Gains							
Mean Gain (\bar{D})		29.4	2.1	2.8	2.8	2.9	40.5
t-statistic		8.5*	5.7*	4.7*	6.4*	5.4*	10.8*
Summary Statement about Objective		MET	MET	MET	MET	MET	MET
*Results of t-test for Dependent Samples Significant at 0.01 Level.							

Oral language, Identification, Part I Reading, Part II Reading, Part III Reading and Total Points. A t-test for dependent samples was applied for N = 44 on each of the six objectives. Results from Table 1 indicate clearly that all program objectives were met.

TABLE 2

GRADE TWO SELF OBSERVATION SCALES DATA (N=44)

	SCALE I Self Acceptance	SCALE II Social Maturity	SCALE III School Affiliation	SCALE IV Self Security
Mean Normal Curve Equivalent (NCE)	50.4	52.5	50.2	50.8
Percentile Rank of NCE Score	50.8	54.8	50.4	51.6
NCE Mean in 1978 (Grade 1)	45.5	40.5	38.1	47.9
Percentile Rank of NCE Mean in 1978	41.5	32.6	28.6	46.0
Mean NCE Gain	+4.9	+12.0	+12.1	+2.9
Percentile Rank Change on NCE Means	+9.3	+22.2	+21.8	+5.6
Summary Statement about SOS Objective*	MET	MET	MET	MET

*An SOS Objective is Met if there is a Positive Mean NCE Gain from the Prior Year.

The first use of NCEs appeared in the analysis of the scores for Grade 2 on the Self Observation Scales (SOS). The SOS instrument is a nationally normed, empirically validated, multi-dimensional instrument in the affective domain for measuring the way children perceive themselves (Stenner and Ratzenmeyer, 1974). The primary level of the SOS measures four different aspects of children's self concept. Each of the four scales is labeled in a positive manner which high scores being most characteristic of the scale factor or name.

Table 2 presents SOS data comparing the Grade 2 class in 1979. This approach seems more feasible in an educational setting in which it is virtually impossible to conduct a true experimental study. The research strategy utilized is a quasi-experimental design in which the class serves as both the control and experimental group. An SOS objective is met if there is a positive mean gain in the NCE score from the prior year. An examination of Table 2 shows that all four SOS objectives (Self Acceptance, Social Maturity, School Affiliation and Self Security) were met, since there was a positive mean NCE gain for each objective. It is especially encouraging to note the magnitude of the gain in Scale II (Social Maturity) and Scale III (School Affiliation). While the NCE mean in 1978 was below the mean on all four scales for the students as first graders, their performance as second graders was above the mean of 50 on all four scales.

In Table 3 are presented the Grade 2 Metropolitan Achievement Test for the Bilingual students (N = 44) both as second and first graders. Comparative data are available in two areas - Reading and Mathematics. If the second grade Normal Curve Equivalent mean is significantly greater than their first grade mean, then that particular achievement test objective is said to be satisfied. It can be seen from an analysis of Table 3 that the Bilingual students showed a significant gain in Mathematics (objective met) but not in Reading (failed to meet objective).

TABLE 3

GRADE TWO METROPOLITAN ACHIEVEMENT TEST COMPARISON (N=44)

	TEST I Reading	TEST II Mathematics
Bilingual Student		
Mean NCE Score (1979)	51.5	49.9
Percentile Rank of Mean NCE 1979 Score	52.9	49.8
Bilingual Student NCE		
Mean in 1978 (Grade 1)	54.5	45.0
Percentile Rank of NCE Mean in 1978 (Grade 1)	58.5	40.6
Mean NCE Gain	-3.0	+4.9
Percentile Rank Change on NCE Mean	-5.6	+9.2
Summary Statement about Metropolitan Achievement Test*	NOT MET	MET

*A Metropolitan Achievement Test Objective is met if there is a positive mean NCE gain from the prior year.

Summary Comments and Observations

Based on an analysis of the data, the following comments and suggestions are made:

1. The Bilingual Project staff should be encouraged as to the extent to which their program objectives were met for Grade 2. All six of the objectives were satisfied.
2. An analysis of Self Observation Scales at the primary level for Grade 2 is very encouraging as all four of the SOS objectives were met.
3. The Metropolitan Achievement Test data for Grade 2 resulted in positive gains in Mathematics but negative gains in Reading.
4. The use of a quasi-experimental design with percentile rank scores changed to Normal Curve Equivalent (NCEs) is recommended for subsequent educational research studies in which the use of randomly assigned control and experimental groups is not possible.
5. In summary, the Bilingual Project staff should be very encouraged with the results for Grade 2 as all six program objectives were met, all four SOS objectives were satisfied and positive gains were detected in one out of the two Metropolitan Achievement tests.

REFERENCES

- Stenner, A.M. and Katzenmeyer, W.G. Self Observation Scales (SOS): Primary Level, National Testing Service, Inc., Durham, North Carolina, 1974.
- Tallmadge, G.K. and Wood, C.T. User's Guide: ESEA Title I Evaluation and Reporting System, RMC Research Corporation, Mountain View, California, 1976.