DOCUMENT RESUME

ED 190 627 TM 800 421

AUTHOR Powell, George D.: Raffeld, Paul C.
TITLE An Investigation of the Equipercentile Assumption and the One-Group Pre/Post Design.
PUB DATE Apr 80
NOTE 27p.: Paper presented at the Annual Meeting of the American Educational Research Association (64th, Boston, MA, April 7-11, 1980).

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Achievement Gains: Achievement Tests: Elementary Education: Elementary School Mathematics: *Evaluation Methods: National Norms: *Pretests Posttests: Reading Achievement: *Research Design: Scores: Standardized Tests
IDENTIFIERS Elementary Secondary Education Act Title I: *Equipercentile Assumption: *RMC Models: SRA Achievement Series

ABSTRACT

The equipercentile assumption states that students in traditional classrooms who receive no other instructional assistance, will maintain their relative rank order over time. To test this assumption, fall to fall test results on the SRA Achievement Tests were obtained for grades 2-3, and 6-7. Total reading and total mathematics growth scale values were used in all analyses. Score distributions were divided into low, medium, and high subgroups of an approximately equal number of matched scores. With the exception of grades 7-8 mathematics scores, the equipercentile assumption was valid only for certain subgroups, not for the total group. These results conflicted with those of a similar study the previous year, providing less than a harmonious picture of the equipercentile assumption. Districts considering Title I evaluation model A should conduct a local study of non-federally impacted children to determine for themselves the appropriateness of the assumption. (CP)

AN INVESTIGATION OF THE EQUIPERCENTILE ASSUMPTION AND THE ONE-GROUP PRE/POST DESIGN

By

GEORGE D. POWELL and PAUL C. RAFFELD

A paper presented at the Annual Meeting of the
American Educational Research Association
Boston, Massachusetts, April 7-11, 1980

POWELL ASSOCIATES, INC.
2734 Jefferson, Suite 205
Austin, Texas 78731

A B S T R A C T

The purpose of this study is to further examine the tenability of the equipercentile assumption as applied to standardized achievement tests. This assumption simply states that standardized achievement tests are so designed that students in traditional classroom instructional settings, who receive no other instructional assistance, will maintain their relative rank order over time.

The apparent value of such an assumption, if valid, is that it is possible, within certain limits, to use the norming group associated with a specific standardized test as a pseudo-control group in evaluation studies. This eliminates the need for a locally-selected control group and offers instead, the use of the pretest mean as the control parameter estimate in a one-group pretest/posttest design.

I N T R O D U C T I O N

With the advent of the 1974 congressional mandate regarding Title I evaluation, RMC Research Corporation was given the task of developing alternative models for Title I evaluation. Among the models chosen by RMC was the one-group pretest/posttest design described by Campbell and Stanley (1962). Tallmadge and Wood (1976) refer to this design as Model A. The new wrinkle, of course, was the introduction of an implied control group and the equi-percentile assumption. It did not take long after the release of the models suggested by RMC for Model A to become the most popular model. Unfortunately, the choice of Model A has been based most often on convenience and ease of implementation, rather than on any sound empirical support for the validity of the model.

To date there has been little documented evidence to support or refute the equipercentile assumption. One

data-based study with reasonable sample sizes was conducted by the present authors in 1979 (Powell, et al, 1979). In this study, pretest and posttest data was obtained from a local school for grades 3, 4, and 5 with sample sizes ranging from 136 to 183. The results of this study suggested some degree of inconsistency, since the equipercentile assumption appeared to hold in general for reading but not for mathematics.

There are two primary drawbacks to this preliminary study. First, the form and level of the test was not constant from pretest to posttest. Since difficulty increases with changes in level, RMC recommended maintaining the same level and form for pretest and posttest when using Model A. It seems reasonable then, to expect a different result when the same level is maintained over a period of time. The second drawback was the omission of the lower and upper ten percent of the distribution. It was assumed that extreme scores (especially low scores) were prone to considerable measurement error and that the regression effect would result in low pretest scores increasing at posttest time by chance alone. However, Tallmadge (1976) clearly notes that if student selection is independent of the pretest, then regression toward the mean will not occur.

The present study examines the equipercentile assumption when following all of the recommended procedures for implementing Model A. Tests are administered at the proper norming date and for purposes of this study, the children are selected because of their lack of involvement in Title I or other federal programs that should influence normal school year growth. The same level of the test is administered at pretest and posttest time.

METHODS

Data were collected from a local school district with an average daily attendance of approximately three thousand children in grades one through twelve. Fall-to-Fall test results were obtained for the 1976-77 school year for grades two to three and for the 1977-78 school year in grades six to seven. No students involved in any federal programs were used in this study.

The total reading and total math growth scale value scores from the SRA Achievement Test were used in all analyses. Children's pretest and posttest scores were matched separately for reading and for mathematics. At grades two to three, there were 104 matched pairs of scores in

both reading and mathematics. At grades 6 to 7, there were 139 matched pairs of scores for both reading and mathematics. It should be noted that there were no special reading or mathematics programs involved; hence, there was no reason to believe that the equipercentile assumption would not hold across the entire range of students.

The range of actual scores obtained and used in the data analysis were between the second and ninety-ninth percentile in reading, and between the sixth and ninety-ninth percentile in mathematics at the second grade level. At the sixth grade level, the range of actual scores were between the fifth and ninety-sixth percentile in reading and between the second and ninety-ninth percentile in mathematics.

In order to determine the extent to which the equi-percentile assumption held throughout the entire test score distribution, the distribution for total reading and total math was divided into thirds. In this way, it was hoped that any systematic deviations in the equipercentile assumption could be isolated according to the initial level of pretest score.

Since growth scale values (GSVs) are not expected

to remain constant from pretest to posttest, the mean growth scale value for pretest and posttest was converted to percentiles and then to Normal Curve Equivalents (NCEs). If the equ percentile assumption held, then the NCE means would be constant from pre to posttest time. A T-test for paired data was used to determine any significant departure from a zero expectation for all sub-groups and total groups.

R E S U L T S

I. End of Grade Two to End of Grade Three Analysis

Figures 1 and 2 represents the frequency distribution of scores in GSV values for both total reading and total mathematics for the end of second grade scores. Based on these pretest distributions, subjects were divided into three categories: low, medium, and high. Categories were determined so that each category would have an approximately equal number of matched scores. Figure 3 shows the summary information for the total reading scores for both pre and posttest. A quick glance at the boxes in Figure 3 indicates that the equipercentile assumption does not appear to be met for the total group or for

the lower two groups; however, for the upper group -- with thirty-six subjects -- there is a very close match of ninety-three percentile for an average on pretest, and a ninety-four percentile for average on posttest. Figure 4 shows similar summary information for total mathematics scores. In this case it is the low group -- with thirty-eight matched subjects -- that shows a very close adherence to the equi-percentile assumption with those thirty-eight children approximating the thirtieth percentile on both pretest and posttest.

Figure 9 presents the necessary information to perform a T-test for matched scores and roughly agrees statistically with the "eyeball" observations above. For the reading achievement between second and third grade, both the low and high sub-groups are not significantly different from a theoretical expectation of the equipercentile assumption. Likewise, in mathematics achievement between the second and third grade, the low and medium groups of stude ts are not statistically different from our equipercentile assumption ($p < .05$). Note that the population value μ pre and μ post are ob-

tained by first obtaining the median for the total group and then each of the three sub-groups on the pretest, and translating those percentile values to equivalent growth scale values using posttest norms.

II. Analysis of Grades Six and Seven Results

Figures 5 and 6 show the frequency distributions for the end of grade six; that is, the pretest for both total reading and total mathematics scores. As in the grade two to three analysis, students' scores were further subdivided into three sub-groups -- low, medium, high -- which had approximately equal and/or matched scores. The range for each sub-group is given in Figures 5 and 6. Summary statistics for total reading at the sixth to seventh grade level is presented in Figure 7. The equipercentile assumption does not seem to hold up well for the total group, or any of the three sub-groups, although it is much closer at the sixth to seventh grade level than it was at the second to third grade level.

Figure 8 presents the results in summary statistics for total mathematics scores. The equipercentile

assumption is a closer approximation for the overall matched group, going from the fifty-eighth to the sixtieth percentile from pre to posttest. It is a good match for the low group, going from the twenty-seventh to twenty-ninth percentile pre to post. It is a good match for the high group, going from the eighty-eighth to the eighty-sixth percentile pre to post. And even for the middle group, there does not seem to be that much difference pre to post as far as an "eyeball" analysis goes. However, information on Figure 9 shows that three of the four comparisons for reading at the sixth to seventh grade level are statistically significant at the ($p < .05$) level. However, none of the differences between obtained and expected growth scale score values on total mathematics scores are significantly different.

III. Discussion

The results of a similar study last year by the authors (Powell et al, 1979), indicated that, using a similar analysis, the equipercentile assumption seemed to hold for both the overall group and the individual low, medium, and high sub-groups for total reading scores for grades

two to three, grades four to five, and over a two-year analysis, grades two to four. At the same time, the equipercentile assumption, in general, did not hold for the total mathematics score for either the total group or the individual sub-groups. Comparing last year's data taken from another school district with this year's data, indeed, provides less than a harmonious picture of the potential of the equipercentile assumption.

The best summary statement that might be culled from both of these studies is that the equi-percentile assumption may or may not be a good assumption, depending not only on the school being investigated but also, the test being used, the test level and form use, and the children taking the test.

A school district interested in using the Title I evaluation Model A would be well-advised to do a similar study of non-federally impacted children in its local school district to see the extent to which it can put its faith in this vital assumption behind Model A, the equipercentile assumption.

Using tests of significance we can find some subgroups, and even some total groups, for which the equipercentile assumption statistically holds true. However, if the equipercentile assumption is to be taken at face value, we find no condition where precisely the same percentile for even small groups of children is maintained from pretest to posttest under those conditions where the equipercentile assumption should obtain; that is, in situations where children are not impacted by other than ongoing school programs.

SUPPLEMENTAL DATA

We have recently obtained information from a large city school district in the southwest that bears further on the problem of the equipercentile assumption. One of the author's concerns of both last year's study and this year's study is that the larger the N size for groups of students, the more closely the equipercentile assumption will hold given that there is a sufficient number of children not impacted by federally-assisted programs. We were able to find approximately 1,500 matched scores for both total reading and total mathematics achievement test scores in a school district using a Spring-to-Spring testing cycle with the ITBS, Level 7, Form 5. Students

were matched on pre and posttest scores and were chosen for their lack of federal program involvement. These children were tested at both the end of grade two and at the end of grade three using the same form and level of the test during the 1978-79 school cycle. Although the data is too recent for the necessary T-tests to have been used, Figure 10 indicates that neither for the total group for reading and mathematics, nor for the quartile of those groups, does the equipercentile assumption visually hold. Contrary to hoping that larger N size will smooth out the variance between the statistical and the visual presentation of the equipercentile assumption, in this one case, it does seem that larger N further dissipates any hope that the equipercentile assumption is a real and consistent phenomenon.

School districts using the Title I Evaluation Model A should be cautioned that prior to implementing the model, a small local study should be done so that they can ascertain for themselves the appropriateness of the equipercentile assumption. Where the equipercentile assumption cannot be met, one of the other two Title I Evaluation Models

might be contemplated, or an alternative model might be presented to USOE based on this and similar types of preliminary analysis.

Work is currently being done by the RMC Research Corporation, under contract with the U. S. Office of Education Office of Evaluation and Dissemination, to use large scale data bases to investigate the validity of the equipercentile assumption. Hopefully, during 1980 more definitive guidelines will be available on the parameters of the equipercentile assumption based on those large-scale data-based studies.

F I G U R E 1

Pre test Distribution of GSV scores for End of Grade 2

TOTAL READING

READING

end Grade 2
reading range = 53 → 311

GSV	FREQUENCY
50 - 65	X
66 - 80	
81 - 95	X X
96 - 110	
111 - 125	X X
126 - 140	X X
141 - 155	X X X X X X X X X X X X
156 - 170	X X X X X X X X X X X X
171 - 185	X X X X X X X X X X
186 - 200	{ X X X X X X X X X X X X X X X X
201 - 215	X X X X X X X X X X X
216 - 230	X X X X X X X X X X X X X
231 - 245	X X X X X
246 - 260	X X X X X X X X X X
261 - 275	X X X X
276 - 290	
291 - 305	
306 - 320	X X X X

READING		N
Low:	53 to 171	35
Medium:	174 to 211	33
High:	216 to 311	36

F I G U R E 2

Pretest Distribution of GSV scores for End of Grade 2
TOTAL MATHEMATICS

MATH

end Grade 2

math range: 108 → 265

GSV	FREQUENCY
50 - 65	.
66 - 80	
81 - 95	
96 - 110	X
111 - 125	X X X
126 - 140	X X X X X X X X X X X X X X X X X X
141 - 155	X X X X X X X X X X X X X X X X X X X X
156 - 170	X X X X X X X X X X X X X X X X X X X X X X X X X X X
171 - 185	X X X X X X X X X X X X X
186 - 200	X X X X X X X X X X X
201 - 215	X X
216 - 230	X
231 - 245	X X X X
246 - 260	
261 - 275	X
276 - 290	
291 - 305	
306 - 320	

MATH

		N
Low:	108 to 151	38
Medium:	153 to 169	33
High:	173 - 265	33

F I G U R E 3

End of Grade 2, End of Grade 3 Summary Statistics
TOTAL READING

Pretest Grade =	2
Year =	1976
Nr tested =	122
Nr matched =	104
% attrition at pretest =	15%

Posttest Grade =	3
Year =	1977
Nr tested =	141
% attrition at posttest =	26%

Form E, Primary I

Pre-post correlation = + 0.72

Form F, Primary I

		Pretest	Posttest
Matched	GSV	196.73 (s=47.81)	259.63 (s=47.03)
	%ile	74%ile	81%ile
Total Tested	GSV	196	258
	%ile	74%ile	80%ile
Low Group (N=35) 2%ile to 58%ile	GSV	147.14 (s=27.32)	219.03 (s=47.71)
	%ile	42%ile	55%ile
Medium Group (N=33) 60%ile to 81%ile	GSV	194.61 (s=10.19)	261.61 (s=23.55)
	%ile	73%ile	81%ile
High Group (N=36) 83%ile to 99%ile	GSV	246.89 (s=29.27)	297.28 (s=26.10)
	%ile	93%ile	94%ile

F I G U R E 4

End of Grade 2, End of Grade 3 Summary Statistics:
TOTAL MATHEMATICS

Pretest Grade	= 2	Posttest Grade	= 3
Year	= 1976	Year	= 1977
Nr tested	= 122	Nr tested	= 141
Nr Matched	= 104	Nr Matched	= 104
% attrition at pretest	= 15%	% attrition at posttest	= 26%

Form E, Primary I

Form F, Primary I

Pre-post correlation = +0.73

		Pretest	Posttest
Matched	GSV	163.86 (s=28.26)	200.47 (s=37.15)
	%ile	59%ile	53%ile
Total Tested	GSV	164	199
	%ile	59%ile	53%ile
Low group (N=38) 6%ile to 44%ile	GSV	137.97 (s=9.93)	176.87 (s=28.13)
	%ile	30%ile	31%ile
Medium group (N=33) 47%ile to 65%ile	GSV	161.00 (s=4.60)	196.09 (s=25.20)
	%ile	56%ile	50%ile
High group (N=33) 69%ile to 99%ile	GSV	196.21 (s-23.40)	232.03 (s=34.39)
	%ile	87%ile	78%ile

F I G U R E 5

Pretest Distribution of GSV scores for End of Grade 6
TOTAL READING

1 9 7 7 S I X T H G R A D E

End Grade 6 — READING

Reading range — 212 → 440

	GSV	FREQUENCY
	200-215	X
	216-230	
	231-245	
	246-260	
	261-275	X X
	276-290	X X X X
	291-305	X X X X X X X X X X
	306-320	X X X X X X X X X X X X X X X
1-St. Dev.	321-335	X X X X X X X X X X X X X
	336-350	X X X X X X X X X X X X X
	351-365	X X X X X X X X X X X X X X X X X X X X X X X X X
+1 St. Dev.	366-380	X X X X X X X X X X X X X X X X X X X X X
	381-395	X X X X X X X X X X X X X X X X X X X X X X X X X
	396-410	X X X X X X
	411-425	X X X X X X X X X X
	425-440	X X X
	441-575	

	READING	N
Low =	212-337	46
Medium =	339-371	40
High =	376-440	52

F I G U R E 6

Pretest Distribution of GSV scores for End of Grade 6

TOTAL MATHEMATICS

1 9 7 7 S I X T H G R A D E

end Grade 6 MATH

Math range 215 → 540

	GSV	FREQUENCY
	200-215	X
	216-230	
	231-245	X
	246-260	
	261-275	X X X
	276-290	X X X X X
	291-305	X X X X X X X X X X X X X
1-St. Dev.	306-320	X X X X X X X X X
	321-335	X X X X X X X X X X X X X X X
	336-350	X X X X X X X X X X X
	351-365	X X X X X X X X X X X X X X X X X X X X X X X X X X X
+1 St. Dev.	366-380	X X X X X X X X X X X X X X X X X
	381-395	X X X X X X X X X X X X X X X
	396-410	X X X X X X X X
	411-425	X X X
	425-440	X X
	441-455	X X X X
	456-470	X X X X
	471-485	X X X X
	486-500	
	501-515	

F I G.U R E 6 (continued)

GSV	FREQUENCY
516-530	X
531-545	X
546-560	
561-575	X X

MATH	N
Low = 215-333	46
Medium = 337-377	51
High = 381-561	42

F I G U R E 7

End of Grade 6, End of Grade 7 Summary Statistics: TOTAL READING

PreTest Grade = 6
Year = 1977
Nr Tested = 169
Nr Matched = 138
% attrition at pretest = 18%
Form F, Green Level

PostTest Grade = 7
Year = 1978
Nr Tested = 165
Nr Matched = 138
% attrition at posttest = 16%
Form E, Green Level

Pre-post correlation = +0.81

		Pretest	Posttest
Matched	GSV	354.84 (s=40.42)	389.12 (s=45.10)
	%ile	60%ile	69%ile
Total Tested	GSV	345	389
	%ile	54%ile	69%ile
Low group (N=46) 5%ile to 48%ile	GSV	308.37 (s=21.94)	347.96 (s=35.35)
	%ile	30%ile	40%ile
Medium group (N=40) .9%ile to 72%ile	GSV	355.93 (s=9.81)	384.40 (s=26.80)
	%ile	61%ile	65%ile
High group (N=52) 75%ile to 96%ile	GSV	395.12 (s=16.98)	429.15 (s=25.13)
	%ile	85%ile	88%ile

F I G U R E 8

End of Grade 6, End of Grade 7 Summary Statistics: TOTAL MATH

PreTest Grade = 6
Year = 1977
Nr Tested = 169
Nr Matched = 139
% attrition at pretest= 18%
Form F, Green Level

PostTest Grade = 7
Year = 1978
Nr tested = 170
% attrition at posttest= 18%
Form E, Green Level

Pre-post correlation = +0.82

		Pretest	Posttest
Matched	GSV	362.32 (s=58.89)	396.69 (s=63.94)
	%ile	58%ile	60%ile
Total Tested	GSV	355	396
	%ile	54%ile	65%ile
Low group (N=46)	GSV	305.28 (s=25.28)	335.93 (s=44.51)
2%ile to 42%ile	%ile	27%ile	29%ile
Medium group (N=51)	GSV	358.24 (s=11.59)	398.43 (s=32.0)
44%ile to 66%ile	%ile	56%ile	61%ile
High group (N=42)	GSV	429.74 (s=49.74)	461.12 (s=44.05)
68%ile to 99%ile	%ile	88%ile	86%ile

F I G U R E 9

SUMMARY STATISTICS FROM LARGE CITY SCHOOL DISTRICT
Grade 2-3, Spring-to-Spring, ITBS, Level 7, Form 5

	Reading		Math	
	TOTAL			
	N = 1525		N = 1494	
	Pre %ile	Post %ile	Pre %ile	Post %ile
	52.36	42.35	29.85	39.85
	Quartile 1			
	N = 379		N = 375	
	Pre %ile	Post %ile	Pre %ile	Post %ile
	12.20	15.11	12.27	14.30
(S.D.)	(9.19)	(14.92)	(12.84)	(16.57)
	Quartile 2			
	N = 377		N = 357	
	Pre %ile	Post %ile	Pre %ile	Post %ile
	46.42	30.98	23.64	29.61
(S.D.)	(7.24)	(19.57)	(15.60)	(23.07)
	Quartile 3			
	N = 426		N = 422	
	Pre %ile	Post %ile	Pre %ile	Post %ile
	67.77	51.85	35.02	48.35
(S.D.)	(5.65)	(20.78)	(19.23)	(25.58)
	Quartile 4			
	N = 343		N = 340	
	Pre %ile	Post %ile	Pre %ile	Post %ile
	86.40	75.70	52.42	69.84
(S.D.)	(5.73)	(18.54)	(19.52)	(22.75)

F I G U R E 1 0

t-test for Matched Scores

	GROUP	$\bar{X}$ PRE	$\bar{X}$ POST	μPRE	μPOST	r PRE-POST	$S_{\bar{x}pp}$	t
	READING							
	Total	197	260	182	234	.72	3.50	+3.14*
2nd	Low	147	219	112	170	.51	7.06	+1.98
↓	Med	195	262	192	245	0	4.54	+3.08*
3rd	High	247	297	263	311	.26	4.95	+0.40
	MATH							
	Total	164	200	186	236	.73	2.50	-5.60*
2nd	Low	138	177	130	167	.47	4.12	+0.49
↓	Med	161	196	161	203	.17	4.39	-1.59
3rd	High	196	232	219	277	.61	4.84	-4.55*
	READING							
	Total	355	389	326	352	.81	2.28	+3.51*
6th	Low	308	348	274	300	.54	4.46	+3.14*
↓	Med	356	384	355	377	.17	4.31	+1.39
7th	High	395	429	408	433	.51	3.08	+2.92*
	MATH							
	Total	362	397	388	422	.82	3.16	+0.32
6th	Low	305	336	274	297	.42	6.10	+1.31
↓	Med	358	398	357	389	.44	4.08	+1.96
7th	High	430	461	471	510	.60	6.60	-1.21

R E F E R E N C E S

Campbell, Donald T., and Stanley, Julian C. Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally College Publishing Company, 1963.

Powell, George D., Schmidt, Jerome R., Raffeld, Paul C. The Equipercentile Assumption as a Pseudo-Control Group Estimate of Gain. Resource Development Institute, Inc., 1979.

Tallmadge, G.K. and Wood, C.T. User's Guide, ESEA Title I Evaluation and Reporting System. RMC Research Corporation, 1976.

Tallmadge, G.K. The Regression Effect, Technical Paper No. 3. RMC Research Corporation, 1976.