DOCUMENT RESUME

ED 336 413 TM 017 179

AUTHOR Schumacker, Randall E.
TITLE Rasch Rating Scale Analysis of Quality Indicators of Elementary and Secondary School Performance.
PUB DATE Mar 91
NOTE 27p.; Paper presented at the Annual Texas Testing Conference (5th, Austin, TX, March 4, 1991).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Administrator Attitudes; Comparative Analysis; *Educational Quality; Elementary Schools; Elementary Secondary Education; Mail Surveys; Outcomes of Education; *Rating Scales; School Districts; School Effectiveness; School Surveys; Secondary Schools; State Surveys; *Superintendents; *Test Construction
IDENTIFIERS *Quality Indicators; *Rasch Model; Texas

ABSTRACT

Types of quality indicators (QIs) for elementary schools and secondary schools in Texas, the selection of indicators by district superintendents in Texas, and the subsequent rating scale analysis using Rasch measurement procedures were studied. QIs were scaled from 1 to 7, with 1 representing "not important", and 7 representing "very important". Superintendents from 1,055 school districts in Texas received two separate 20-item QI rating scales for elementary schools and secondary schools. After a follow-up, the response rate was 71.5% (N=754). Results indicate that both the elementary school and secondary school rating scales were reliable. Item calibrations defined a continuum of important/not important variables, with item separation better defined for elementary school QIs as compared to secondary school QIs. The Rasch measurement model permitted comparisons of QI logits or standard scores (normal curve equivalents); consequently, it may have practical value for comparing similar schools. Eleven data tables and a 13-item list of references are included. An appendix provides descriptions of selected terms. (SLD)

Rasch Rating Scale Analysis

of

Quality Indicators of Elementary and Secondary School Performance

Randall E. Schumacker, Ph.D.

University of North Texas

Paper Presented
at
FIFTH ANNUAL TEXAS TESTING CONFERENCE
March 4, 1991
Stouffer Hotel, Frio Room, 1 - 2 pm
Austin, Texas

ABSTRACT

Quality indicators of school performance have received national and state attention. Texas has recently proposed and will be implementing an Academic Excellence Indicator System. The selection of demographic variables to group similar school campuses and, more importantly, the selection of outcome measures for comparison purposes is therefore of paramount importance. The present study reports findings on types of indicators for elementary and secondary schools, the selection of indicators by district superintendents in Texas, and the subsequent rating scale analysis using Rasch measurement procedures. The Rasch rating scale analysis measurement model may have practical utility in comparing similar schools.

Rasch Rating Scale Analysis

of

Quality Indicators of Elementary and Secondary School Performance

Introduction

Quality indicators of school performance have received national and state attention. Recent passage of Senate Bill 1 in the State of Texas mandates the implementation of an Academic Excellence Indicator System beginning in January, 1991. The selection of variables to group similar school campuses and the adoption of certain outcome variables is therefore of paramount importance.

Identifying the domains that need to be monitored is the first step in developing indicators and the next step involves defining what indicators are available and how they should be distinquished from other data (Murname & Raizen, 1988, pp 27-39). Oakes (1986) additionally emphasized that for a measure to be used as an indicator it must have a reference point so that a judgment can be made as to whether the condition being described is getting better or worse. Smith (1988) added that educational indicators should either assess or be related to desired outcomes of the educational system. Basically, multiple measures are preferred to single measures of school performance.

Selection of Educational Indicators

In 1988, the President of the United States emphasized educational issues by holding an educational summit meeting of governors. Shavelson, McDonnell, and Oakes (1989) in response suggested that state educational indicator systems should include information in the following areas: (a) **inputs** (fiscal, material, teacher quality, student backgrounds, etc.); (b) **process** (school context, organization, curriculum, teaching/instructional quality, etc.); and (c) **outcomes** (student achievement, participation, attitudes, aspirations, etc.). Odden (1990) further suggested that single indicators not be used and that an indicator system should provide measures from the various components of the educational system to examine how they work together and whether any changes in the system exist over time. Other research indicated that monitoring outcomes alone would not provide enough information about changes over time and might mislead comparisons (Guskey & Kifer, 1989). Schumacker & Brookshire (1990) subsequently identified quantitative and qualitative indicator variables for each input, process, and outcome component of a statewide indicator system.

Methods and Procedures

The present research reports findings on types of quality indicators for elementary and secondary schools, the selection of indicators by district superintendents in Texas, and the subsequent rating scale analysis using Rasch measurement procedures.

Subjects

Superintendents were choosen to respond to two separate quality indicator rating scales because of their unique administrative position within independent school districts in Texas. All of the superintendents from 1,055 districts in the State of Texas were mailed a cover letter, business reply envelope, and two rating scales. There are 1,076 districts in Texas, but special districts were excluded from the survey. A follow-up was conducted three weeks after the initial mailing with a 71.5 % response rate. The district superintendent's response rate was determined to be adequate and equally representative of the various regions as indicated in Table 1 (chi-square goodness of fit = 5.97, df = 19, p = .998).

Insert Table 1 Here

Instrument

Two separate twenty item quality indicator rating scales were developed for elementary and secondary schools after a review of quality indicators (National Association of Elementary School Principals, 1984 & 1990; Schumacker & Brookshire, 1990). The quality indicators were scaled from 1 to 7 (1 = not important to 7 = very important) and are presented in Tables 2 and 3, respectively. The two separate rating scales were reviewed by area district administrators prior to a state-wide survey. The internal consistency reliability of the instruments were r = .89 (elementary quality indicator scale) and r = .90 (secondary quality indicator scale).

Insert Table 2 Here

Insert Table 3 Here

Factor analysis of the twenty quality indicators for each elementary and secondary school rating scale was also conducted (Schumacker & Brookshire, 1991). Factor analysis indicated how the quality indicators as multiple measures would group together. For elementary school indicators, Factor 1, the factor accounting for

the largest proportion of variance, was related to the "curriculum" of the school. Factor 2 was related to "school professional environment"; Factor 3 was related to "school statistics" or less academic measures; Factor 4, which shared several quality indicators with the other factors, was related to the more objective measures of "school quality"; and Factor 5 was related to the "teacher and instruction". For secondary schools, Factor 1 was related to "academic/vocational choice". Factor 2 related to the more objective measures of "school statistics". Factor 3 was related to the "national measures of achievement". Factor 4 was an "activity" factor. The most important groupings of quality indicators were indicated on Factor 1 as "curriculum" indicators for elementary schools and "academic/vocational choice" indicators for secondary schools. No unidimensional outcome however was obtained for either elementary or secondary school performance.

Rasch Analysis

The Rasch measurement model unlike the classical true score model provides the following advantages in scaling and interpretation:

1. linear scale ("logit" metric)
2. sample-free item calibrations
3. scale-free person measures
4. common scale for persons and items

Rasch rating scale analysis utilizes a "step" interpretation to the analysis of Likert-style questionnaires (Andrich, 1978a; 1978b). The relative difficulties of the "steps" in a rating scale of items is usually accompanied by a fixed set of points common to all items. This expectation is incorporated into two components:

$$\delta_{ik} = \delta_i + \tau_k$$

The scale value of an item and the location of the k'th step in the item are indicated relative to the item's scale value.

The rating scale model which is the probability of person "n" responding in a category "x" to an item "i", then becomes (Wright & Masters, 1982):

$$\pi_{nix} = \frac{\exp \sum_{j=0}^{x} [\beta_n - (\delta_i + \tau_j)]}{\sum_{k=0}^{m} \exp \sum_{j=0}^{k} [\beta_n - (\delta_i + \tau_j)]}$$

The Rasch rating scale analysis model will produce separate person and item parameters. Person parameters will be conditioned out of item calibration, enabling sample-free calibration; and item parameters will be conditioned out of person measurement, enabling scale-free measurement. This will entail reporting "integer scoring" values from Likert-type attitude scales. The Rasch measurement model utilizes a logistic item operating curve with common slope.

Research Questions

The research questions to be investigated for the elementary and secondary rating scales are:

1. Where is item i located? (item calibration)
2. How precise is this calibration? (error of calibration)
3. Where is person n located? (person calibration)
4. How precise is this measure? (error of measurement)
5. How well do both these responses fit the measurement model? (item and person fit: separation and reliability)

Results

Elementary and secondary school quality indicators were analyzed separately for both item calibration and person calibration. A glossary of terms is provided in the Appendix.

Elementary

Rasch analysis item calibrations for the elementary school quality indicators are in Table 4. The logit values were rescaled into normal curve equivalent (NCE) scores which range from 1 to 99 with a mean of 50 and standard deviation of 21. Because NCE scores have characteristics of both percentile ranks and equal interval scale, they can be used to compare indicators within or between either rating scale similar to the logits. Table 5 indicates the summary statistics for both the logit and NCE scores.

The item separation index assesses the extent to which items (or persons) are separated along a common scale. The ITEM SEP value indicates that the ADJ S.D. is 16.95 times greater then the RMSE value which indicates a good separation of items along the important/not important rating scale continuum. The reliability of this item separation is 1.00. Table 6 presents the calibrated step summary information whereby the STEP CALIBR. values indicate the distance between response categories.

Table 7 indicates the summary statistics for both the logit and NCE scores based on person (superintendent) calibration. The item separation index indicates the extent to which superintendents

were separated along the important/not important continuum (or extent of individual differences). The reliability of .88 compares favorably with the internal consistency reliability of .89 indicated earlier.

Insert Tables 4, 5, 6, 7 Here

Secondary

Rasch analysis item calibrations for the secondary school quality indicators are in Table 8. The logit values were also rescaled into normal curve equivalent (NCE) scores. Table 9 indicates the summary statistics for both the logit and NCE scores.

The ITEM SEP value indicates that the ADJ S.D. is 8.89 times greater then the RMSE value which indicates a good separation of items along the important/not important rating scale continuum. The reliability of this item separation is .99. Table 10 presents the calibrated step summary information whereby the STEP CALIBR. values indicate the distance between response categories.

Table 11 indicates the summary statistics for both the logit and NCE scores based on person (superintendent) calibration. The item separation index indicates that superintendents were not as separated along the important/not important continuum (or showed less individual differences). The reliability of .91 compares favorably with the internal consistency reliability of .90 indicated earlier.

Insert Tables 8, 9, 10, 11 Here

Conclusions and Discussion

Results indicate that both the elementary and secondary rating scales are reliable. Item calibrations defined an important/not important variable continuum with item separation better defined for elementary, as compared to, secondary quality indicators. Rasch analysis provided the added benefit of providing:

1. linear scale ("logit" metric)
2. sample-free item calibrations
3. scale-free person measures
4. common scale for persons and items

The Rasch measurement model therefore permits comparisons of quality indicator logits or standard scores (NCE's) either between or within each rating scale.

Recommendations

A statewide educational system is comprised of three components (inputs, process, outcomes). It is therefore recommended that a model be developed which links the components and evaluates their relationship over time. This basically involves combining input, process, and outcome variables into an explanatory model. Also, since districts and their schools have great diversity in these three educational system components, it is recommended that school profiles be developed which compare similar schools on outcome measures that monitor changes over time. The Rasch rating scale

analysis and measurement model is useful because it provides a technique for comparing and monitoring performance of indicators over time and from school to school. It essentially provides a common scale whereby these comparisons can be made. It is postulated that this model may have practical utility in comparing schools on outcome measures as well.

References

Andrich, D. (1978a). Scaling attitude items constructed and scored in the Likert tradition. Educational and Psychological Measurement, 38, 665-680.

Andrich, D. (1978b). Application of a psychometric rating model to ordered categories which are scored with successive integers. Applied Psychological Measurement, 2, 581-594.

Guskey, T.R. & E.W. Kifer (1989). Ranking School Districts on the Basis of Statewide Test Results: Is it meaningful or misleading? Paper presented at the American Educational Research Association annual meeting, San Francisco, California.

Murname, R.J. & Raizen, S.A. (1988). Improving Indicators of the Quality of Science and Mathematics Education in Grades K-12. National Academy Press, Washington, D.C. 1-231.

National Association of Elementary School Principals (1984). Standards for Quality Elementary Schools: Kindergarten through Eighth Grade. NAESP: Virginia, 1-60.

National Association of Elementary School Principals (1990). Standards for Quality Elementary & Middle Schools: Kindergarten through Eighth Grade, Revised Edition. NAESP: Virginia, 1-58.

Oakes, J. (1986). Educational Indicators: A guide for policy-makers. Rutgers University Center for Policy Research in Education. New Brunswick, NJ.

Odden, A. (1990). Educational Indicators in the United States: The Need for Analysis. Educational Researcher, 19(5), 24-29.

Schumacker, R.E. & W.K. Brookshire (1990). Defining Quality Indicators. Texas Study of Secondary Education Research Journal, 46, 31-34.

Schumacker, R.E. & W.K. Brookshire (1991). Factor Analysis of Quality Indicators for Elementary and Secondary Schools. Texas Researcher, 1, 81-89.

Shavelson, R.J., McDonnell, L.M., & J. Oakes (1989). Indicators for monitoring mathematics and science education(Eds). The Rand Corporation: Santa Monica, CA.

Smith, M. (1988). Educational Indicators. Phi Delta Kappan, 69(7), 487-491.

Wright, B.D. & G.N. Masters (1982). Rating Scale Analysis: Rasch Measurement. MESA PRESS: Illinois.

Table 1: Superintendent Response by Regional Service Center

Regional Service Center	N	n
I.	37	28
II.	44	33
III.	41	29
IV.	55	43
V.	29	19
VI.	57	36
VII.	99	71
VIII.	48	30
IX.	40	26
X.	80	62
XI.	76	58
XII.	79	62
XIII.	56	42
XIV.	47	29
XV.	45	25
XVI.	68	51
XVII.	61	41
XVIII.	33	24
XIX.	13	9
XX.	47	36
Total	1055	754 (71.5 %)

Table 2: Elementary Quality Indicator Rating Scale

ELEMENTARY SCHOOLS

QUALITY INDICATOR RATING SCALE

Below is a list of characteristics that have been suggested as quality indicators of elementary schools. These characteristics are probably not of equal importance when judging the quality of a school. Please indicate the importance you would place on each of these characteristics when evaluating elementary schools in your district.

QUALITY INDICATOR CHARACTERISTICS	PLEASE CIRCLE NOT IMPORTANT						VERY IMPORTANT
1. State mandated test results	1	2	3	4	5	6	7
2. School goals and objectives	1	2	3	4	5	6	7
3. Student attendance statistics	1	2	3	4	5	6	7
4. Staff training and background	1	2	3	4	5	6	7
5. Principal's role in school	1	2	3	4	5	6	7
6. Promotion statistics	1	2	3	4	5	6	7
7. Approved curriculum content	1	2	3	4	5	6	7
8. Budget allocations	1	2	3	4	5	6	7
9. Extracurricular activities	1	2	3	4	5	6	7
10. Student attitude information	1	2	3	4	5	6	7
11. Teacher evaluation criteria	1	2	3	4	5	6	7
12. Program evaluation reports	1	2	3	4	5	6	7
13. School environment	1	2	3	4	5	6	7
14. School/Community/Business relation	1	2	3	4	5	6	7
15. Inservice training programs	1	2	3	4	5	6	7
16. Crime rates in schools	1	2	3	4	5	6	7
17. Curriculum implementation	1	2	3	4	5	6	7
18. Teacher attitude information	1	2	3	4	5	6	7
19. Instruction yields desired outcomes	1	2	3	4	5	6	7
20. Curriculum content and breadth	1	2	3	4	5	6	7

Table 3: Secondary Quality Indicator Rating Scale

SECONDARY SCHOOLS

QUALITY INDICATOR RATING SCALE

Below is a list of characteristics that have been suggested as quality indicators of secondary schools. These characteristics are probably not of equal importance when judging the quality of a school. Please indicate the importance you would place on each of these characteristics when evaluating secondary schools in your district.

QUALITY INDICATOR CHARACTERISTICS	PLEASE CIRCLE NOT IMPORTANT						VERY IMPORTANT
1. State mandated test results	1	2	3	4	5	6	7
2. Dropout rates	1	2	3	4	5	6	7
3. Student attendance statistics	1	2	3	4	5	6	7
4. ACT/SAT results	1	2	3	4	5	6	7
5. Grade-point average statistics	1	2	3	4	5	6	7
6. Promotion statistics	1	2	3	4	5	6	7
7. Follow-up status after high school	1	2	3	4	5	6	7
8. Suspension and expulsion statistics	1	2	3	4	5	6	7
9. Extracurricular activities	1	2	3	4	5	6	7
10. Outside activities	1	2	3	4	5	6	7
11. Student attitude information	1	2	3	4	5	6	7
12. National merit semi-finalist	1	2	3	4	5	6	7
13. Advanced placement credit	1	2	3	4	5	6	7
14. Computer literacy	1	2	3	4	5	6	7
15. Gains in special education	1	2	3	4	5	6	7
16. Vocational choices	1	2	3	4	5	6	7
17. College choices	1	2	3	4	5	6	7
18. Job placement	1	2	3	4	5	6	7
19. College placement	1	2	3	4	5	6	7
20. Crime rates in schools	1	2	3	4	5	6	7

Table 4: Elementary School Quality Indicators:[a] Item Calibrations

Quality Indicator	Logit	NCE[a]
PRINCIPALS ROLE IN SCHOOL	1.74	86.54
SCHOOL ENVIRONMENT	1.30	77.30
INSTRUCTION YIELDS OUTCOMES	.74	65.54
SCHOOL GOALS AND OBJECTIVES	.54	61.34
TEACHER ATTITUDE INFORMATION	.41	58.61
STAFF TRAINING/BACKGROUND	.35	57.35
CURRICULUM CONTENT/BREADTH	.31	56.51
CURRICULUM IMPLEMENTATION	.23	54.83
STUDENT ATTENDENCE STATISTICS	.02	50.42
APPROVED CURRICULUM CONTENT	-.03	49.37
PROGRAM EVALUATION REPORTS	-.18	46.22
SCHOOL/COMMUNITY/BUSINESS	-.25	44.75
INSERVICE TRAINING PROGRAMS	-.25	44.75
BUDGET ALLOCATIONS	-.29	43.91
STUDENT ATTITUDE INFORMATION	-.36	42.44
TEACHER EVALUATION CRITERA	-.39	41.81
STATE MANDATED TEST RESULTS	-.77	33.83
PROMOTION STATISTICS	-.80	33.20
CRIME RATES IN SCHOOLS	-.89	31.31
EXTRACURRICULAR ACTIVITIES	-1.43	19.97

[a] NCE = 50 + 21 (Logit)

Table 5: Elementary School Quality Indicators: Scale Summary Statistics[a]

Scale	N	M	S	Min	Max
Logit	20	0.0	.75	-1.43	1.74
NCE	20	50.0	15.68	19.97	86.54

[a] Rasch analysis (RMSE = .04; ADJ.S.D. = .73; ITEM SEP = 16.95; ITEM SEP REL. = 1.00)

Table 6: Elementary School Quality Indicators: Calibrated Step Summary

CATEGORY LABEL	STEP VALUE	OBSERVED COUNT	STEP CALIBR.	STEP ERROR	EXPECTED SCORE CALIBRATIONS STEP-.5	AT STEP	STEP+.5
1	0	128	NONE			EXTREME	-1.98
2	1	205	-.98	.09	-1.98	-1.46	-1.13
3	2	498	-1.02	.06	-1.13	-.84	-.57
4	3	1706	-.97	.04	-.57	-.27	.08
5	4	3502	-.02	.03	.08	.50	1.01
6	5	4637	.91	.02	1.01	1.62	2.49
7	6	3963	1.99	.02	2.49	EXTREME	

Table 7: Elementary School Quality Indicators:[a] Superintendents

Scale	N	M	S	Min	Max
Logit	733	1.34	.78	-1.61	5.17
Items	733	91.80	12.20	20.00	140.00

[a] Rasch analysis (RMSE = .27; ADJ.S.D.= .74; PERSON SEP = 2.76; PERSON SEP REL. = .88)

Table 8: Secondary School Quality Indicators:[a] Item Calibrations

Quality Indicator	Logit	NCE[a]
STUDENT ATTENDENCE STATISTICS	.81	67.01
STUDENT ATTITUDE INFORMATION	.56	61.76
DROPOUT RATES	.37	57.77
COLLEGE PLACEMENT	.32	56.72
COMPUTER LITERACY	.30	56.30
ACT/SAT RESULTS	.12	52.52
COLLEGE CHOICES	.10	52.10
CRIME RATES IN SCHOOLS	.08	51.68
JOB PLACEMENT	.02	50.42
STATE MANDATED TEST RESULTS	-.04	49.16
FOLLOW-UP AFTER HIGH-SCHOOL	-.07	48.53
PROMOTION STATISTICS	-.08	48.32
VOCATIONAL CHOICES	-.09	48.11
GAINS IN SPECIAL EDUCATION	-.13	47.27
GRADE-POINT AVERAGE STATISTICS	-.22	45.38
ADVANCED PLACEMENT CREDIT	-.23	45.17
EXTRACURRICULAR ACTIVITIES	-.24	44.96
SUSPENSION/EXPULSION STATISTIC	-.42	41.18
NATIONAL MERIT SEMI-FINALIST	-.44	40.76
OUTSIDE ACTIVITIES	-.70	35.30

a

NCE = 50 + 21 (Logit)

Table 9: Secondary School Quality Indicators: Scale Summary Statistics[a]

Scale	N	M	S	Min	Max
Logit	20	0.0	.35	-.70	.81
NCE	20	50.0	7.44	35.30	67.01

[a] Rasch analysis (RMSE = .04; ADJ.S.D. = .34; ITEM SEP = 8.89; ITEM SEP REL. = .99)

Table 10: Secondary School Quality Indicators: Calibrated Step Summary

CATEGORY LABEL	STEP VALUE	OBSERVED COUNT	STEP CALIBR.	STEP ERROR	EXPECTED SCORE CALIBRATIONS STEP-.5	AT STEP	STEP+.5
1	0	152	NONE			EXTREME	-2.14
2	1	337	-1.22	.09	-2.14	-1.58	-1.22
3	2	782	-.96	.05	-1.22	-.92	-.63
4	3	2676	-1.06	.03	-.63	-.30	.08
5	4	4616	-.05	.02	.08	.55	1.11
6	5	3977	1.04	.02	1.11	1.78	2.71
7	6	3963	1.99	.02	2.71	EXTREME	

Table 11: Secondary School Quality Indicators: Superintendents[a]

Scale	N	M	S	Min	Max
Logit	717	.81	.81	-1.42	5.32
Items	717	82.10	13.90	20.00	140.00

[a] Rasch analysis (RMSE = .24; ADJ.S.D.= .77; PERSON SEP = 3.15; PERSON SEP REL. = .91)

APPENDIX

a

GLOSSARY OF TERMS

TERM	DESCRIPTION
Normal Curve Equivalent	Transformed score with $\underline{M}$ = 50; $\underline{S}$ = 21
RMSE	Root mean square standard error
ADJ S.D.	Estimate of "true" standard deviation
ITEM (or PERSON) SEP	Coefficient of separation (ADJ S.D./RMSE)
ITEM (or PERSON) REL.	Coefficient of separation reliability (Rasch equivalent to KR-20/Cronbach Alpha) Estimate of the ratio of "true" variance to the "observed" variance: $(\text{ADJ S.D.}/\text{S.D. of MEASURE})^2$

[a] Wright, B.D., J.M. Linacre, M.S. Schulz 1990). A User's Guide to BIGSCALE. Chicago, Illinois: Mesa Press, 33-34.