DOCUMENT RESUME

ED 398 253 TM 025 202

AUTHOR	Baker, Sheldon R.; And Others
TITLE	Towards the Nonstochastic Recalibration of the Teacher Made Test and Project Scoring Protocols as a Measure of Student Performance Magnitudes: The Instructional Effectiveness Coefficient for Group and Individual "Empiricism without Experimentation."
PUB DATE	95
NOTE	11p.; Version of a paper presented at ECER (Bath, England, United Kingdom, 1995).
PUB TYPE	Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Algorithms; *Educational Assessment; Educational Research; Elementary Secondary Education; Epistemology; *Instructional Effectiveness; Models; Performance Factors; Psychometrics; Qualitative Research; *Research Design; Research Methodology; *Scoring; *Teacher Made Tests; Test Construction
IDENTIFIERS	*Calibration; Edumetric Properties of Tests; Empirical Research; *Stochastic Analysis

ABSTRACT

A paradigm for the recalibration of teacher-made assessment that assesses and evaluates in one operation is formulated. The effort to make the classroom the primary source of educational research activity is contingent on redefining educational research as empirical and not experimental. This emphasizes that the empirical analysis of instructional effectiveness is a question of design effectiveness and calls for a new methodology for quantitative analysis, called edumetrics. Edumetrics consists of the formulation of a descriptive index coefficient, which can describe magnitude of student performance using teacher-made instruments as the numerical base. The Instructional Effectiveness Coefficient is discussed as an edumetric device that amalgamates the assessment and evaluation of instructional effectiveness and empirical educational research. Two edumetric algorithms are presented to assess and evaluate teacher-made tests. A new research epistemology is formulated that is quantitatively intuitive and empirical. The formulated model is nonstochastic in its logic. (SLD)

TOWARDS THE NONSTOCHASTIC RECALIBRATION OF THE TEACHER MADE TEST AND PROJECT SCORING PROTOCOLS AS A MEASURE OF STUDENT PERFORMANCE MAGNITUDES: THE INSTRUCTIONAL EFFECTIVENESS COEFFICIENT FOR GROUP AND INDIVIDUAL "EMPIRICISM WITHOUT EXPERIMENTATION"

By

Sheldon R. Baker, Ed.D., Ohio County Schools
John Paterson, Ed.D., West Virginia University
H. Lawrence Jones, Ed.D., Superintendent - Ohio Schools
Bonnie Ritz, Ed.D., Ohio County Schools
Patricia Pockl, M.A., Ohio County Schools

Business Address: Ohio County Board of Education
2203 National Road
Wheeling, WV 26003
Phone: (304) 243-0300; Fax: (304) 243-0328

Contact Person: Sheldon R. Baker, Ed.D.
782 Des Moines Avenue
Morgantown, WV 26505
Phone: (304) 292-3896

Original paper given at ECER, Bath, England 1995 as "Experimentation Without Control: A New Formulation for Classroom Research"

Abstract

A paradigm for the recalibration of teacher made assessment is formulated which assesses and evaluates in one operation. The Instructional Effectiveness Coefficient is discussed as an edumetric device which amalgamates the assessment and evaluation of instructional effectiveness and empirical educational research. A new research epistemology is formulated which is quantitatively intuitive and empirical. The model which is formulated is nonstochastic in its logic.

Introduction

The effort to make the classroom as the prime source of educational research activity is contingent upon redefining educational research as empirical and not experimental. Generic research designs which employ true or quasi type of control or contrasting groups have given rise to complex multivariate designs which have yet to be proven decisive as yielding casual models which can lead to more effective classroom instruction (Draper 1995).

By defining classroom research as empirical, we are emphasizing that the empirical analysis of instructional effectiveness is a question of design effectiveness and, therefore, can only be studied in the classroom at the time the instructional design is being executed and managed to its completion.

Strategic Assumptions

I. The basic question for classroom research is how effective is an instructional design?

II. Since instructional effectiveness is the basic research question, parametric statistic and inference is not an appropriate paradigm for the analysis of student test or project performance.

III. In light of assumptions I and II, a new methodology for quantitative analysis needs to be developed. This is called Edumetrics.

IV. Edumetrics consist of the formulation of a descriptive index coefficient, which can describe magnitude of student performance using teacher made instruments as the numerical base.

V. Classroom research is empirical and not experimental. Experimentation involves a contrasting or control group. Since classroom research is non-stochastic, the researcher's only interest is to determine how effectively an instructional design works. Empiricism requires an uncontrolled intervention which is instructional design.

VI. The IEC is offered as an edumetric index coefficient, which can assess and evaluate instructional effectiveness in the same operation, thus eliminating the need for a control group.

VII. The IEC can be employed with any edumetric method of tests or qualitative project calibration for the assessment of quality magnitudes.

The following assumptions, which are strategic and modular, set our philosophic case for a new epistemology of data description which is empirically descriptive and edumetric. We shall depart from customary experimental design and challenge the basic assumption upon which traditional assumptions are made with regard to the experimental approach to classroom research. What we shall propose is heretical and intuitive. As with any new quantitative algorithm, this paper begins its formulation on the frontier of a new discipline which we shall call Edumetrics, a concept which has rarely, if ever, been considered as educations' field of measurement as psychometrics is to psychology.

For too many years traditional educational experimental research has been dominated by variants of the Campbell Stanley model of experimental design (Campbell have speciously relegated teacher made efforts at assessment and evaluation of student performance to the level of routine drudgery. The use of traditional experimental design and inferential statistics has obscured the value of teacher made assessments in determining the effectiveness of an instructional design. In light of the aforementioned situation, we shall recommend the use of a newly developed edumetric algorithm, which will both assess and evaluate student performance either as a group or an individual function in such a manner as to assess and evaluate in one operation. The algorithm, as we shall describe it in the paper, will permit classroom research without a contrast or contrasting group through the use of a new descriptive statistic.

Tactical Considerations

We shall show how the IEC-G and the IEC-I can be used to recalibrate teacher made assessments so that instructional effectiveness can be inferred in either da summative or formative mode with either selected or constructed response instruments or with performance magnitudes of the summative attributes of qualitative projects.

Many thousands of hours of teacher time are being wasted in assessment of teacher made tests and projects because their efforts are not being utilized as basic summative and formative indices of instructional effectiveness in experimental design. The classroom teacher research question is, "How effective is my instructional design in producing adequate student productivity?"

Since the basic question of the classroom is one of instructional effectiveness, questions of statistical probalistic inference are not applicable. The classroom teacher is also not interested in establishing laws or generalizing to a population. The question of also not interested in establishing laws or generalizing to a population. The question of instructional effectiveness is pragmatic and limited in its explanatory timeline.

Therefore a new assessment procedure which is edumetric and descriptive is needed in order to transform teacher efforts to assess and evaluate student test and project performances.

We shall, therefore, present two edumetric algorithms which assess and evaluate in one operation using teacher made extended and constructed response tests and project points.

The algorithms are simple to calculate and interpret.

The Instructional Effectiveness Coefficient
For Groups and Individual Performance
(IEC-G and IEC-I)

Steps in Calculation of the IEC-G:

1. Class size: minimum of 15 students, maximum class size sample
2. Calculate the average for a group of scores = $\overline{X}_D$
3. Calculate the average of the scores below the mean = $\overline{X}_L$
4. Calculate the average score or mean $\overline{X}_H$ for scores above the group mean or $\overline{X}_D$.
5. Subtract $\overline{X}_L$ and $\overline{X}_H$ from the distribution mean and square each result or

   $(\overline{X}_D - \overline{X}_H)^2 = A$

   and

   $(\overline{X}_D - \overline{X}_L)^2 = B$
6. Divide $\frac{A}{B}$ and take the square root = $\sqrt{\frac{A}{B}}$ = IEC-G

or

$$\sqrt{\frac{(\overline{X}_D - \overline{X}_H)^2}{(X_D - X_L)^2}} = \text{IEC-G}$$

Interpretive Standards

1.26 - 1.51 = Exceptional

1.10 - 1.25 = Above the average

.90 - 1.0 = Average

.60 - .84 = Below the average

Less than .60 = Deficient. Needs intervention for improvement.

A preliminary validity study indicates that the IEC-G correlates .63 with a measure of negative skewness. The relationship suggests only minimum overlap between a measure skewness and IEC-G.

Steps in Calculation of IEC-I:

1. Calculate the average group

2. Divide each score by the distribution average or mean = $\overline{X}_D$

3. Multiply each X_I by IEC-G

   or $[(X_I) * (\text{IEC-G})] = \text{IECI}$

4. The student's score is a handwith of

$$\left(\frac{\overline{X}_I}{\overline{X}_D}\right) \text{ to } \left(\frac{\overline{X}_I}{\overline{X}_D}\right) * (\text{IEC-G})$$

   A. If score contains any value from 9.0 to 1.10 then the performance is average or acceptable.

   B. If score contains any value 1.11 to 1.25 then performance is above the average.

C. If score contains any value from 1.26 to 1.51, then performance is exceptional.

D. If the score is between .75 and .89, performance is below the average.

E. If score is less than .74, performance is deficient.

Reprise

In this conceptual paper, we have explicated the logic and use of a new descriptive quantitative algorithm, which recalibrates the teacher made test and project, thereby permitting descriptive statements of group and individual performance without the use of a control or contrasting group. The key question which the IEC-G and IEC-I is designed to answer is the question of instructional effectiveness at the point the assessments are made. The summative instrument of a project or test can become formative in allowing for remedial or developmental intervention. The IEC-G and IEC-I free the teacher from the dominance of those specialists who view classroom research as a scientific enterprise which is encumbered by complex statistical analysis which denigrates the teacher made assessments.

Implication Toward the Development in Edumetric Epistemology:

The use of the IEC-G opens up research in instructional effectiveness, as well as forming a unique discipline of educational assessment called Edumetrics, a descriptive and intuitive recalibration of teacher made efforts to assess and evaluate aggregate and individual student performance. As a new technology of assessment, the IEC-G-I can be calculated with a simple bivariate statistical calculator, or it can be adapted for spreadsheet data description and analysis without the imposition of complex multivariate statistical designs which are contingent for use upon the assumption of stochastic educational realities. The IEC-G appears to have its own domain of measurement and it may, upon deeper mathematic analysis, become the fifth statistical moment.

The IEG can be used to analyze a semester of summative tests in the following manner:

Each series of test scores is converted to a percentage * 10 score. This yields a score range from 1-9.9 across the entire series of teacher made tests for a semester. The IEC-G can be employed to analyze and evaluate the semester overall performance of a class and thereby providing the teacher a quality assurance index of student test performance and a part of a total class evaluation program.

Other ratio types of indices can be developed which are non stochastic and descriptive of student performance. This we think should be the focus of Edumetrics as a discipline of education. As we see it now, Edumetrics bypasses traditional error mathematics in focusing on scaler and vector factors, which are implicit within the manner in which we have employed Euclidian distances as the gradient of the mean.

References

Campbell, D.T. & Stanley, J.C. (1963). Experimental Design for Research and Teaching. Chicago: Rand McNally.

Cook, T.D. & Campbell, D.T. (1979). Quasi Experimentation in Design and Analysis Issues for Field Settings.

Cooper, H.M. (1989). "Statistical Versus Traditional Procedures for Summarizing Research Review." Psychological Bulletin. 87, 442-449.

Draper, D. (1995). "Inference and Hierarchical Modeling in the Social Sciences." Journal of Education and Behavioral Statistics. 20, 115-147.