DOCUMENT RESUME

ED 062 403 TM 001 344

AUTHOR Greene, John F.; Zirkel, Perry A.
TITLE Scoring Creativity Tests by Computer Simulation: A Validation of Prediction Equations.
PUB DATE Apr 72
NOTE 7p.; Paper presented at the annual meeting of the American Educational Research Association (Chicago, Illinois, April 1972)

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Computers; Correlation; Creativity; *Creativity Tests; Models; Multiple Regression Analysis; Predictive Ability (Testing); *Predictive Validity; Rating Scales; Scoring; Test Reliability; *Test Scoring Machines; *Test Validity
IDENTIFIERS *Torrance Test of Creative Thinking

ABSTRACT

The general usefulness of selected predictions equations for computer simulated scoring of creativity tests was studied. This was carried out by testing previously established prediction equations for samples drawn from similar populations. (CK)

SCORING CREATIVITY TESTS BY COMPUTER SIMULATION: A VALIDATION OF PREDICTION EQUATIONS

John F. Greene
University of Bridgeport

Perry A. Zirkel
University of Hartford

Paper Presented at the Annual Convention of the American Educational Research Association, Chicago, April 1972

Scoring Creativity Tests by Computer Simulation: A Validation of Prediction Equations

John F. Greene[1]
University of Bridgeport

Perry A. Zirkel
University of Hartford

In recent years several studies have attempted to utilize the computer to simulate human behavior, especially human rating behavior (e.g. Page and Paulus, 1968; Archambault, 1969; Greene, 1970; Whalen, 1970). These attempts included scoring essays and creativity tests by computer. In each study, the researcher developed a multiple regression prediction equation and then proceeded to either empirically or statistically cross-validate the equation within a randomly selected partition of the original sample. None of the prediction equations, however, have actually been employed and then evaluated in samples drawn from other similar populations. Consequently, the objective of this study was to determine the general usefulness of selected prediction equations.

Greene (1970) developed successful computer prediction models for activities four through seven of the Torrance Tests of Creative Thinking, Verbal-Form A, (TTCT) (Torrance, 1966). Activities four through seven are Product Improvement (toy elephant), Unusual Uses (cardboard boxes), Unusual Questions (cardboard boxes), and Just Suppose (if clouds had strings, what would happen?) respectively. Each activity is scored for three dimensions of creativity: fluency, flexibility, and originality. A flexibility score, however, is not

---

[1]This research paper extends the principle author's doctoral dissertation. The study was initiated while the author was a USOE research fellow at The University of Connecticut. Appreciation is expressed to Dieter Paulus and Joseph Renzulli for assistance.

determined for the sixth activity, Unusual Questions. In Greene's (1970) study four judges rated the responses of 153 subjects. Analysis of variance procedures were used to provide an estimate of the pooled ratings of the judges, as suggested by Winer, (1962, pp. 124-132). A stepwise multiple regression technique was emplcyed to maximize the prediction of each subject's scores for each activity. The predictors included the actuarial and dictionary parameters generated earlier by computer. Besides the full model, restricted and forced regression models were generated. The entire computerized scoring procedure was then evaluated in a cross-validation sample.

METHOD

In the present study, the TTCT was administered to 190 students in 2 central Connecticut suburban schools. These students were in grades 4-7 and represented a sample similar to the New York sample used in the original investigation. The responses of each subject were independently rated by four judges. Adjusted reliability estimates of these pooled ratings were determined by analysis of variance procedures. The prediction equations developed by Greene (1970) were then used to evaluate the student responses. In order to determine the usefulness of the prediction equations, the computer-generated scores for the full, restricted, and forced models were correlated with the human ratings. These correlations were attenuated for the unreliable aspects of the criterion variable. Shrinkage was determined by contrasting the correlation coefficients with those reputed for the cross-validation sample in the original study.

RESULTS/CONCLUSIONS

The adjusted pooled reliability estimates are presented in Table I.

TABLE I

ADJUSTED POOLED RELIABILITY ESTIMATES FOR FOUR JUDGES USING ANALYSIS OF VARIANCE

	Dimension		
Activity	Fluency	Flexibility	Originality
4	.97	.95	.86
5	.98	.96	.90
6	.98	---	.82
7	.98	.89	.81

**All correlations significant at .01 level.

These estimates ranged from .97 to .98 for fluency and .89 to .96 for the flexibility ratings. The originality coefficients, ranging from .81 to .90 for the four activities, were noticeably higher than those found in the original study.

The attenuated correlations between the computer generated scores and the human ratings for the full, restricted, and forced models of the validation sample as well as the results of the original study are given in Table II.

TABLE II

RESULTS OF ORIGINAL
AND NEW SAMPLES:
ALL MODELS

Activity	Dimension	Full Model R	Full Model r attenuated orig. sample	Full Model r attenuated new sample	Restricted Model R	Restricted Model r attenuated orig. sample	Restricted Model r attenuated new sample	Forced Model R	Forced Model r attenuated orig. sample	Forced Model r attenuated new sample
4	Fluency	.99	.96	.88	.99	.96	.86			
4	Flexibility	.91	.81	.69	.90	.86	.67			
4	Originality	.84	.79	.56	.84	.83	.55			
5	Fluency	.96	.93	.88	.95	.96	.94			
5	Flexibility	.85	.89	.68	.84	.91	.70			
5	Originality	.87	.90	.74	.86	.92	.78			
6	Fluency	.97	.95	.90	.97	.96	.91			
6	Originality	.80	.84	.57	.79	.86	.59			
7	Fluency	.92	.87	.86	.90	.85	.84			
7	Flexibility	.84	.56	.72	.83	.59	.73	.73	.77	.77
7	Originality	.73	.49	.47	.70	.64	.48	.60	.70	.52

**All correlations are significant at .01 level

The four full model correlations in the new validation sample for fluency exceeded .86. Shrinkage was limited to .08. Lower correlations, ranging from .68 to .72, and greater shrinkage, up to .21, were obtained for the flexibility dimension. Although shrinkage was noted, the full model prediction equations for fluency and flexibility were deemed stable and useful.

The originality coefficients were bound by .42 and .70, with shrinkage ranging from .08 to .40. Except for Activity five, these results indicate that the original full model prediction equations for originality are relatively unstable.

The results of the restricted and forced models paralleled those of the full model. Although higher coefficients were generally found, appreciable differences in predictive ability were not realized. Consequently, the usefulness of these models is restricted to parsimonious considerations.

While computer simulation models continue to appeal to the fancy of man's mind, the usefulness of these models after initial development must be continually considered. Otherwise, such models will be considered an mere academic games.

BIBLIOGRAPHY

Archambault, F. A Computerized Approach to Scoring Verbal Responses To The Torrance Tests of Creative Thinking. Unpublished Doctoral Dissertation, The University of Connecticut, 1969.

Greene, J.F. Scoring Creativity Tests by Computer Simulation. Unpublished Doctoral Dissertation, The University of Connecticut, 1970.

Page, E. and Paulus, D. The Analysis of Essays by Computer. USOE Final Report, 1968.

Torrance, E. Torrance Tests of Creative Thinking. Princeton, New Jersey: Personnel Press, 1966.

Whalen, T. E. A Computer Analysis of Mechanical Proficiency and Overall Quality in Seventh Grade English Essays. Unpublished Doctoral Dissertation, The University of Connecticut, 1970.

Winer, B. Statistical Principles in Experimental Design. New York: McGraw Hill, 1962.