PREDICTION OF SUCCESS ON THE STATE PROFICIENCY TEST FROM PERFORMANCE ON THE METROPOLITAN ACHIEVEMENT TEST

BY

FELTON CHARLES DAVIS

PREDICTION OF SUCCESS ON THE STATE PROFICIENCY TEST FROM PERFORMANCE ON THE METROPOLITAN ACHIEVEMENT TEST

A Research Paper

Presented to

the Graduate Council

Austin Peay State University

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts

---

by

Felton Charles Davis

March, 1979

To the Graduate Council:

I am submitting herewith a Research Paper written by Felton Charles Davis entitled "Prediction of Success on the State Proficiency Test From Performance on the Metropolitan Achievement Test." I recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Arts, with a major in Psychology.

Garland E. Blair

Major Professor

Accepted for the
Graduate Council:

William H. Ellis

Dean of the Graduate School

ACKNOWLEDGEMENTS

The author wishes to express sincere appreciation to Dr. Garland E. Blair, Professor of Psychology, Austin Peay State University, who assisted and counseled him during the course of the study.

Appreciation is extended to Mr. Moorehead, of the Montgomery County School System, who made it possible for the author to obtain the data used in this study.

Very special appreciation is extended to the author's wife, Honey, and to their children and grandchildren for their patience, assistance, and support, without which this study would not have been completed.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

Introduction to the Problem

Recently, considerable attention has been directed toward State Proficiency Tests (SPT). Floyd (1978) states that probably no single topic regarding education has been so much in the news in the past year as proficiency testing. Thirty-five of the fifty states now have some form of this testing. According to Pipho (1978), Arizona, Florida, Kentucky, and Maryland have tied grade promotion to minimal competency standards. Proficiency testing, minimal competency testing, state-wide assessment of minimal competency, minimal achievement level testing, or survival skills testing, all refer to the back-to-the-basics issue and all are referred to in this paper as State Proficiency Tests.

Pipho says that it is difficult to find two states which have taken an identical approach to Proficiency Testing. Savage (1978) quotes the Executive Director of the Oregon Education Association as being initially enthused with the idea, but says it has now turned into a disaster, with a terrible hodge-podge of standards across the state, so that a student can't transfer from one district to another without considerable difficulty. Oregon has

let each of its school districts establish and measure its own set of competency standards.

Glass (1978) refers to Florida's program, which has a minimum standard (cutoff score) of 70 percent set by the state, as bizarre, and says that the educational system is in the throes of disruption and dislocation. Floyd states that Tennessee has mandated a test as a prerequisite to a high school diploma. This policy becomes effective in the 1981-82 school year.

McClung (1978) states that many of the programs requiring students to pass a minimal proficiency or competency test as a prerequisite to a high school diploma are designed and/or implemented in an inequitable manner, and some may be illegal due to any one of the following reasons: inadequate phase-in periods, inadequate match between test and instruction, racial discrimination, and model program provisions. McClung's reasons will be discussed more fully later in this paper.

Wise (1978) says that there appears to be two different problems that minimum competency testing is designed to solve. The first is the devaluation of the high school diploma. Some recipients of a high school diploma lack the basic skills of reading, writing, and arithmetic. The second is a more profound problem:

a minority of students fail to acquire the basic skills; a minority of teachers fail to teach the basic skills. Wise does not believe that minimum competency testing will improve education. He says that reason and evidence do not justify the belief that minimum competency testing will help poor students to learn or poor teachers to teach.

Bracey (1978) believes that minimum competency testing will not alleviate the problems, but will exacerbate them. Glass predicted that the Florida minimal competency testing would be suspended before the years end, 1978. At this writing, 1979, it has not. Proficiency testing appears firmly entrenched, disastrous or not. According to Hart (1978), starting in July 1980, California school districts are not to award a high school diploma to any student who has not met the locally adopted proficiency standards.

A look at two programs that have some similarity--cutoff scores of 70 percent--Florida, already declared a disaster by Glass, and Tennessee's where the failure rate is also high, will outline the problem.

Glass reports that Florida's results of 35 percent failure rate in mathematics and a 10 percent failure rate in reading/writing caught much of the public and

the students by surprise. Floyd reports much the same surprise in Tennessee where statewide results showed that the following percentages of students failed to attain the required levels of proficiency: Language Arts, 37; Mathematics, 53; Spelling 54; and, Reading, 58. These surprisingly high failure rates could be due not to student performance, or lack of it, but to an arbitrarily set and unrealistic minimum standard, 70 percent.

Statement of the Problem

The purpose of this paper is to determine the relationship between the SPT administered to the eighth-grade students in Montgomery County, Tennessee, School System, in the Spring of 1978, and the MAT administered to the same students when they were in the sixth-grade. Predictive instruments for success on the SPT will be constructed based on this relationship and a rational minimal standard will be recommended.

Review of the Literature

It appears that one of the chief difficulties in the effective use of criterion-referenced tests is the establishment of a meaningful but useful minimum standard. Glass, after receiving information on how Florida set their minimum standard, remarked that they had committed

every wrong against setting sensible and safe standards. According to Glass, the following dialogue tells how Florida's minimum standards were set:

> Question: "How did you finally decide what score was passing?"
>
> Answer: "Well, we worked with people at Florida State and Iowa City, and we spent a lot of time in workshops with local curriculum people. Of course, it's always going to be subjective; but we had to have political credibility in the state. We finally defined 'mastery' as 70% of the items correct. We found that the vast majority of the schools in Florida regard 70% as passing. Of course it's somewhat subjective."

Brickell (1978), in answer to "how high the minimum?", gives this example:

> Recently a group of high school teachers made two minimum competency tests for the end of tenth-grade. One in English and one in mathematics. Any student who failed would be given remedial work, possibly two years of it and possibly no diploma. Good reason for the teachers to make the test fairly easy and good reason for the students to try fairly hard. The tests were of fifth-grade difficulty with a passing score of 60% required. Not very hard. About 25% of the tenth-graders flunked the English test and about 50% flunked the mathematics test. Brickell asked the teachers and principals afterward, that since the students had not learned it in five years, they may not learn it in two more. If remediation failed, then what? How many diplomas can you withhold at commencement? The answer was, maybe 3%. This implies that 97% of the students must pass the competency. He then asked the teachers and principals: what can you guarantee--not wish--that 97% of all graduates

> can do? They answered first-grade work, maybe second-grade work.

It is evident that the setting of minimum standards is not any easy task. Fremer (1978), in response to Glass' article and in support of proficiency testing, thinks that the move to minimum standards can be a positive step. But Fremer recommends use of L. Negelsky's method and William Angoff's variation of it to set the minimum standard. This method presents judges with the unfamiliar task of imagining a minimally acceptable student and guessing the probability that the student will answer a given question correctly. Glass considers the method wrong because he says its rationale, collecting the judgments before the score reports are distributed, is conspicuously bad.

Ebel (1978) also supports proficiency testing, but expresses concern over the difficulty of setting minimum standards. He believes that the test should be a domain-sampling test and not an objective-referenced test; that it should be designed to indicate a pupil's general level of achievement, not the particular competencies the pupil may or may not have learned. He offers an example to set minimum standards:

> "One example might be a 100-item multiple choice test on which the average score for all pupils would be about 65 and the minimum passing score

> about 40. Making the test considerably easier (average score for all pupils 80, minimum passing score 60) would result in better accuracy for the pass-fail decisions at the cost of less accuracy in measuring excellence. The decision on which of these alternatives or some other option is chosen will probably be based on judgments of the relative importance of accuracy in measurement at each end of the scale. The important point here is that measurement should be provided at both ends."

McClung states that proficiency testing could be unfair due to an inadequate match between test and instruction. This concept should be considered in terms of both curricular validity and instructional validity. Fairness requires a school's curriculum and instruction to be matched in some way with whatever is later measured by the test. In other words the test would be unfair if it measured what the school never taught. Curricular validity is a measure of how well test items represent the objectives of the curriculum. An analysis of curricular validity would require a comparison of the test objectives with the school's course objectives. A competency test should also have instructional validity, which is an actual measure of whether the schools are providing students with the instruction in the knowledge and skills measured by the test.

McClung further states that content validity, as

defined by the American Psychological Association, does not insure either curricular or instructional validity. The concepts are related but distinguishable. Content validity is a measure of how well test items represent the performance domain that the test purports to measure. But it is not necessarily a measure of how well the test items and performance domain represent either a particular school's curricular objectives or instruction received. Questions of instructional and curricular validity are relevant, given either of two general purposes of competency testing: 1) to measure the student's mastery of the school's curriculum, or 2) to predict the minimum competence required in the adult world.

McClung states that while substantial numbers of white middle-class students cannot meet minimum competency standards, there is some evidence that a disproportianate percentage of black and hispanic students will be adversely affected by the competency test requirements. While not opposed to competency testing per se, some black parents in desegregated communities see a racial motive behind competency testing. They say that competency testing was not a major concern at either

black or white schools until the schools in their district were desegregated, at which time competency testing was introduced ostensibly "to protect standards." The effect can be resegregation within the schools according to test results (or other forms of tracking) since unequal educational opportunities may cause black children to score lower than their white counterparts.

Hart (1978), a former teacher, and the individual responsible for drafting California's minimum competency program, states that any legislation to require proficiency testing, particularly if it includes sanctions against those who fail, should build in protections so that students alone do not bear the burdens of skill mastery. The California experience indicates that protections should include community involvement in standard setting, periodic assessment, parent notification, and appropriate instruction over time. The responsibility for standard setting was delegated to local school districts.

CHAPTER II

METHOD

Subjects

The sample was composed of 685 students from the Montgomery County, Tennessee, School System, who were administered the SPT in the eighth-grade, and for whom all of the subtest scores for the SPT and the MAT (administered in the sixth-grade) were available.

Procedures

Zero order correlations were obtained between each of the following subtests for the SPT and the MAT: SPT subtests, Mathematics (MA); Spelling (SP); Language (LG); and, Reading (RD); MAT subtests, Word Knowledge (WK); Reading (R); Total Reading (TR); Language (L); Spelling (S); Mathematics Computation (MX); Mathematics Concepts (MC); Mathematics Problem Solving (MP); Mathematics Total (MT). Means (M) and Standard Deviations (SD) were also computed for the 15 variables.

These data were then utilized to compute multiple correlations between the SPT and the MAT. Regression equations and standard errors of the estimate (SE) were computed between the SPT and the relevant subtests of the MAT.

The regression equations were used to predict a score on the SPT from each stanine, 1 through 9, which could be achieved on the relevant subtests of the MAT. In this manner the probability of passing the SPT, achieving 70 percent, was computed. All calculations were done with a Processor technology Sol-20 micro-computer.

CHAPTER III

RESULTS

The zero order correlations among all subtests are presented in Table 1. All subtests on each of the test batteries are highly correlated with each other. In fact, the various subtests of the MAT correlate just as highly with the subtests of the SPT as they correlate with each other. The high correlation between the relevant subtests: MAT MT to SPT MA = .800; MAT S to SPT SP = .695; MAT L to SPT LG = .752; and, MAT TR to SPT RD = .793, made it possible to make useful predictions without having to resort to more complicated multiple regression techniques. Multiple correlations computed between the SPT and the MAT did not result in a useful increase in variability.

Tables 3 through 6, compiled by using the regression formulae, indicate the predicted SPT scores from each stanine achievable on the MAT. From the tables it is evident that a student who is considered average by national standards, a stanine of 5, has only the following chances of achieving the minimum standard on the SPT: Mathematics, 41%; Spelling, 38%; Language, 64%; and, Reading 42%. Table 2 indicates that 40% of the students are below the 5th stanine.

CHAPTER IV

DISCUSSION

State-wide results are pertinent for only the states concerned. However, the data presented herein indicate the probable results both, of state retention of these minimum standards, and if these standards are applied on a national level. Students considered only slightly less than average, stanine of 4, have less than a 75% chance of passing the minimum standards on three of the four tests. Table 2 indicates that 23% of the students are below the 4th stanine.

It is expected that some students will not acquire the basic skills prior to graduation. But, it is doubtful that any state would want to fail more than 40% of its students, especially that proportion considered average by national standards. However, that is what is statistically evident if the present minimum standard prevails.

If the performance of students on national norms is to be considered inadequate then it must be concluded that the entire educational system is in dire need of revision. In such case, the students should not be further penalized for being the victims of an inadequate education.

If the penalizing of students is to be prevented,

it is suggested that the predicted SPT scores from the MAT 3d stanine, to include the Standard error, be accepted as the minimum standard for each test. Tables 3 through 6 indicate these scores to be MA 55; SP 54; LG 59; and, RD 51. Table 2 indicates that below the 3d stanine we can expect only 11% of the students. By accepting scores that fall within -1SE of this minimum standard, a failure rate of less than 11% can be expected.

More research is needed with the MAT as a predictive instrument for success on the SPT. Caution is advised when trying to generalize the results of this study, which was limited to one county in Tennessee.

REFERENCES

Bracey, Gerald W. Some reservations about competency testing. Phi Delta Kappan, 1978, Vol. 59, 549-552.

Brickell, Henry M. Seven key notes on minimum competency testing. Phi Delta Kappan, 1978, Vol. 59, 589-591.

Ebel, Robert L. The case for minimum competency testing. Phi Delta Kappan, 1978, Vol. 59, 546-549.

Floyd, Otis. Proficiency testing - carrying the basics forward. Tennessee Teacher, Vol. 46, 14-15.

Fremer, John. In response to Gene Glass. Phi Delta Kappan, 1978, Vol. 59, 605-606.

Glass, Gene V. Minimum competence and incompetence in Florida. Phi Delta Kappan, 1978, Vol. 59, 602-605.

Hart, Gary K. The California pupil proficiency law as viewed by its author. Phi Delta Kappan, 1978, Vol. 59, 592-595.

McClung, Mark S. Are competency testing programs fair? legal? Phi Delta Kappan, 1978, Vol. 60, 397-400.

Pipho, Chris. Minimal competency standards. Todays Education, 1978, Vol. 67, 34-37.

Savage, David G. Minimum competencies - the Oregon approach. Educational Leadership, Vol. 36, 12-15.

Wise, Arthur E. Minimum competency testing: Another case of hyperrationalization. Phi Delta Kappan, 1978, Vol. 59, 596-598.

APPENDIX: TABLES

Table 1 shows the zero-order correlation between the SPT and the MAT and also the correlation within the SPT and MAT plus the means and standard deviations. N = 685

Table 1

	MA	SP	LG	RD	WK	R	TR	L	S	MC	MCC	MSP	MT	SC	SS	MEAN	SD
MA	XXXX	.651	.749	.759	.646	.633	.661	.722	.610	.764	.761	.740	.800	.624	.664	64.14	18.189
SP		XXXX	.698	.699	.642	.603	.645	.676	.695	.603	.610	.600	.642	.593	.639	65.19	17.515
LG			XXXX	.861	.725	.670	.725	.752	.664	.663	.712	.691	.732	.697	.730	69.78	20.751
RD				XXXX	.775	.757	.793	.765	.658	.655	.735	.725	.745	.756	.769	61.99	21.184
WK					XXXX	.838	.843	.798	.719	.647	.734	.729	.748	.813	.838	4.17	1.810
R						XXXX	.942	.793	.675	.645	.729	.727	.743	.799	.807	4.47	2.061
TR							XXXX	.823	.725	.663	.757	.747	.768	.835	.853	4.35	1.978
L								XXXX	.768	.764	.810	.794	.838	.826	.825	4.35	1.936
S									XXXX	.675	.685	.682	.729	.690	.722	4.86	1.998
MX										XXXX	.803	.800	.918	.679	.706	4.13	1.926
MC											XXXX	.841	.924	.772	.786	4.20	1.803
MP												XXXX	.924	.749	.771	4.28	2.049
MT													XXXX	.781	.804	4.24	1.973
SC														XXXX	.858	4.48	2.123
SS															XXXX	4.42	1.800

Table 2

Stanines and Percentages

Stanines	1	2	*3	4	5	6	7	8	9
Cumulative Percentages	4	11	23	40	60	77	89	96	100
Percent in each Stanine	4	7	12	17	20	17	12	7	4

*3 = stanine proposed for minimum standard

Table 2 indicates what percentage of students can be expected to be in each stanine, and the cumulative percentages of students per stanine.

Table 3

Predicted Scores on the SPT

Mathematics

MAT Stanine	Predicted SPT Score	Z-Score	Probability of Achieving 70%
1	40.215	-2.732	.003
2	47.595	-2.055	.0197
*3	54.975	-1.378	.0838
4	62.355	-0.701	.2420
**5	69.735	-0.024	.4052
6	77.115	.652	.7422
7	84.495	1.329	.9082
8	91.875	2.006	.9778
9	99.255	2.683	.9963

Standard Error = 10.904

*3 = suggested minimum standard

**5 = average student

The figures in Table 3 indicate the predicted SPT scores from the MAT stanines. The z-scores indicate how far the predicted score is from the minimum standard.

Table 4

Predicted Scores on the SPT

Spelling

MAT Stanine	Predicted SPT Score	Z-Score	Probability of Achieving 70%
1	41.714	-2.245	.0071
2	47.804	-1.762	.0392
*3	53.894	-1.278	.1003
4	59.984	-0.795	.2119
**5	66.074	-0.312	.3783
6	72.164	.172	.5675
7	78.254	.655	.7454
8	84.344	1.138	.8729
9	90.434	1.621	.9474

Standard error = 12.599

*3 = the suggested minimum standard
**5 = the average student

The figures in Table 4 indicate the predicted SPT scores from the MAT stanines. The Z-scores indicate how far the predicted score is from the minimum standard.

Table 5

Predicted Scores on the SPT

Language

MAT Stanine	Predicted SPT Score	Z-Score	Probability of Achieving Minimum Standard (70%)
1	42.722	-1.989	.0233
2	50.829	-1.401	.0808
*3	58.888	- .812	.2090
4	66.946	- .223	.4129
**5	75.004	.366	.6443
6	83.062	.954	.8289
7	91.119	1.543	.9382
8	99.178	2.132	.9834
9	107.236	2.720	.9965

Standard error = 13.687

*3 = the suggested minimum standard
**5 = the average student

The figures in Table 5 indicate the predicted SPT scores from the MAT stanines. The z-scores indicate how far the predicted score is from the minimum standard.

Table 6

Predicted Scores on the SPT

Reading

MAT Stanine	Predicted SPT Score	Z-Score	Probability of Achieving Minimum Standard (70%)
1	33.571	-2.822	.0024
2	42.065	-2.164	.0154
*3	50.559	-1.506	.0655
4	59.053	- .848	.2005
**5	67.547	- .190	.4247
6	76.040	.468	.6808
7	84.535	1.126	.8708
8	93.029	1.783	.9625
9	101.523	2.442	.9927

Standard error = 12.908

*3 = the suggested minimum standard
**5 = the average student

The figures in Table 6 indicate the predicted SPT scores from the MAT stanines. The z-scores indicate how far the predicted score is from the minimum standard.