DOCUMENT RESUME

ED 201 673 TM 810 270

AUTHOR Powers, Stephen; Sabers, Darrell
TITLE An Investigation of Ethnic Group Differences in Testwiseness at the Third, Fifth, and Seventh Grade.
SPONS AGENCY National Inst. of Education (ED), Washington, D.C.
PUB DATE Apr 81
GRANT NIE-G-80-0076
NOTE 22p.; Paper presented at the Annual Meeting of the American Educational Research Association (65th, Los Angeles, CA, April 13-17, 1981).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS American Indians; Anglo Americans; Blacks; Elementary Education; *Ethnic Groups; *Grade 3; *Grade 5; *Grade 7; Hispanic Americans; Multiple Regression Analysis; Reading Achievement; Scores; *Sex Differences; *Test Wiseness
IDENTIFIERS *Tucson Public Schools AZ

ABSTRACT

A test measuring testwiseness was administered to 811 pupils in grades 3, 5, and 7 in the Tucson Unified School District. The pupils represented four ethnic groups: American Indian, Anglo, Black, and Hispanic. By controlling for reading ability using scores on a standardized reading test, the California Achievement Tests (CAT) or Comprehensive Tests of Basic Skills (CTBS), as a covariate, significant sex-by-ethnicity interactions were found. The use of reading comprehension as a covariate resulted in findings substantially different from previous research or from what may be concluded from looking at obtained testwiseness scores as a dependent variable. Due to small numbers of pupils in some groups and inconsistencies across grades, specific group results were regarded as extremely tentative. (Author/RL)

A B S T R A C T

A test measuring testwiseness was administered to 811 pupils in the third, fifth, or seventh grades in the Tucson Unified School District. The pupils represented four ethnic groups: American Indian, Anglo, Black, and Hispanic. By controlling for reading ability using scores on a standardized reading test (CAT or CTBS) as a covariate, we were able to find significant sex-by-ethnicity interactions. The use of reading comprehension as a covariate resulted in findings substantially different from previous research or from what we would conclude from looking at obtained testwiseness scores as a dependent variable. Due to small numbers of pupils in some groups and inconsistencies across grades, we must regard specific group results as extremely tentative.

An Investigation of Ethnic Group Differences in Testwiseness at the Third, Fifth and Seventh Grade

Introduction

Testwiseness (TW) was first identified as a possible effector of reliability by Thorndike (1951). A viable conceptual framework for the construct was provided by Millman, Bishop and Ebel (1965) whose definition of TW is employed in this study:

> a subject's capacity to utilize the characteristics and formats of the test and/or the test-taking situation to receive a high score.

Researchers have shown that TW can be measured (Gibb, 1964; Millman, 1966; Slakter et al., 1970a; Woodley, 1973; Sagtelmit, 1978), that TW can be taught (Gaines and Jongsma, 1974; Sarmento, 1979), and that training in TW skills improves performance on tests designed to measure TW (Gibb, 1964; Slakter et al., 1970a; Moore, Schutz, and Baker 1966; Langer et al., 1973). More relevant to the purposes of this study, researchers have shown that instruction in TW raises the scores of students measured on standardized tests such as the Comprehensive Tests of Basic Skills (Gaines and Jongsma, 1974), the Stanford Reading Test (Callemmon, 1973), the Metropolitan Readiness Test (Oakland, 1972).

The purposes of this study were (1) to compare the testwiseness of four ethnic groups--American-Indians, Black, Hispanic and Anglo, (2) to compare the testwiseness of male and female students and (3) to investigate any interaction between ethnicity and sex at the third, fifth, and seventh grade levels. Because previous research has found significant positive correlations between TW and verbal ability (Diamond and Evans, 1972), reading ability was used as a covariate.

Method

Sample. Thirty-nine teachers and 1,128 students at the third, fifth and seventh grade levels of the Tucson Unified School District participated in this study. Five junior high schools and thirteen elementary schools were represented in this study. Schools were selected in order to obtain representative ethnic groups and to obtain representative geographical areas of Tucson. The initial group was reduced to 811 when only those students with complete data sets were included in the analysis. The ethnicity and sex of the participants are presented in Table 1.

Table 1. Participants in the Testwiseness Study

	Third		Fifth		Seventh		
Ethnicity	M	F	M	F	M	F	Total
American Indian	4	4	4	1	7	10	30
Black	12	9	6	3	5	13	48
Hispanic	42	35	70	71	95	81	394
Anglo	22	36	51	57	91	82	339
TOTAL	80	84	131	132	198	186	811

Instrumentation. Four areas of TW were selected from Millman et al., 1965. These areas had been used in previous TW research (Crehan et al., 1978; Crehan et al., 1974; Slakter et al., 1970a; Slakter et al., 1970b). As stated by Millman et al.:

> The examinee should be able to
>
> 1. Select the option which resembles an aspect of the stem.
> 2. Eliminate options which are known to be incorrect and to choose from among the remaining options.
> 3. Eliminate similar options; i.e., options which imply the correctness of each other.
> 4. Eliminate those options which include specific determiners.

The general approach to the construction of a TW test was adapted from Slakter. Items 1-10 contained five items which measured risktaking. The remainder of the test (items 11-26) measured testwiseness subdivided into four subscales: (1) Stem Option Resemblance (items 11, 15, 19, 23), (2) Absurd Options (items 12, 16, 20, 24), (3) Similar Options items 13, 17, 21, 25), and (4) Specific Determiners (items 14, 18, 22, 26). New items were written to be relevant for third through seventh grade students. Items were written with a simple vocabulary to produce a TW Test which would be more valid for the majority of students who were below average in reading ability.

Procedure. TW tests were administered to participants in third, fifth and seventh grades by their teachers who volunteered to participate in this study. Teachers previously received two hours of inservice about testwiseness principles which included procedures for administering the TW test.

Analysis. The following variables were analyzed in this study:

1. Reading Comprehension was measured by the California Achievement Test, Book 13D, Reading Comprehension subtest (with third grade participants), the Comprehensive Tests of Basic Skills, Form S, Level 2, Total Reading Test (with fifth grade participants), and the California Achievement Test, Book 17C, Reading Test (with seventh grade participants). The fifth and seventh grade reading tests had been administered as part of a districtwide testing program in the fall of 1980.

2. Risktaking and Testwiseness as measured by the Standard Educational Test developed by the authors and previously described in this paper.

3. Sex (Male, Female)

4. Ethnicity (American-Indian, Black, Hispanic, Anglo)

Data were analyzed using a multiple regression approach with three dummy variables representing ethnicity and one dummy variable representing sex. Three variables were created to represent the ethnic by sex interaction term.

Reading scores were entered first in the equation followed by ethnicity, sex and ethnicity x sex interaction terms (for more information on this approach, refer to Kerlinger and Pedhazur, 1973, Chapters 6-11).

Results

Alpha coefficients (Cronbach, 1951) were computed for each subscale of the TW test as well as for the total test. Refer to Table 2.

Table 2. Reliability (Alpha) Coefficients for Subscales of the Test-wiseness Test

Grade	Risk	Stem Option	Absurd Option	Similar Option	Specific Determiner	Total Scale*
3	.88	.20	.32	.02	.08	.38
5	.70	.37	.46	.06	.12	.54
7	.39	.24	.21	.18	.08	.41

*Total scale is items 11-26.

Slakter et al. (1970a, p. 120) reported for his test of testwiseness the following median reliability for the same items administered across Grades 5-11: Stem-option subscale, .31; absurd-options subscale, .25; specific-determiners subscale, .08, and the similar-options subscale, .46. The median reliability of the total TW measure was .44.

Crehan et al. (1978, page 42) reported KR-20 reliabilities in a cross-sectional analysis for total TW score for fifth grade .40 and .41 and for seventh grade .44 and .51. We agree with Crehan et al. (1978) that the reliabilities are not high, but for a scale of 16 items, they were adequate for detecting group differences.

A multiple regression analysis was computed for third grade students on the first step of each regression analysis, reading scores were entered to control for that variable. Ethnic groups did not differ significantly, F (1,159) = 1.18. Sex differences in testwiseness did occur, F (1,161) = 5.66 p = .019. Of special interest in this research was an ethnicity by sex interaction, F (3,155) = 2.70, p = .048. Refer to Tables 3-5. A graph of the interaction is presented in Figure 1. Reading means for third, fifth and seventh grades are implemented in Appendix .

Table 3. Regression of Testwiseness Scores on Ethnicity, Sex and E x S Interaction with Reading Scores as Covariate for 164 Third Grade Students

Source	R^2	df	F[a]	p
Covariate	.197	1,162	39.86	.000
Ethnicity	.215	3,159	1.18	.319
Ethnicity + Sex	.249	1,158	7.15	.008
Sex	.225	1,161	5.66	.019
Sex + Ethnicity	.249	3,158	1.70	.169
Sex + Ethnicity + E x S Interaction[b]	.286	3,155	2.70	.048

[a]Each F ratio is computed only for the last variable entered into the regression equation.

[b]Tests for covariate by ethnicity interaction, F (3,156) = .63 and covariate by sex interaction, F (1,160) = 1.04 were nonsignificant.

Table 4. Testwiseness Means Adjusted for Reading Scores for 164 Third Grade Students[a]

Ethnicity	Male	Female	Total
American Indian	8.86	9.06	8.96
Black	10.16	7.85	9.17
Hispanic	9.04	8.94	8.99
Anglo	10.40	9.09	9.59
TOTAL	9.57	8.89	9.23

[a]The regression weight (b weight) used in the adjustment process was .1560.

Table 5. Testwiseness Obtained Score Means for 164 Third Grade Pupils

Ethnicity	Male	Female	Total
American Indian	7.50	8.25	7.88
Black	10.25	7.33	9.00
Hispanic	8.93	9.14	9.02
Anglo	10.13	9.50	9.76
TOTAL	9.40	9.06	9.23

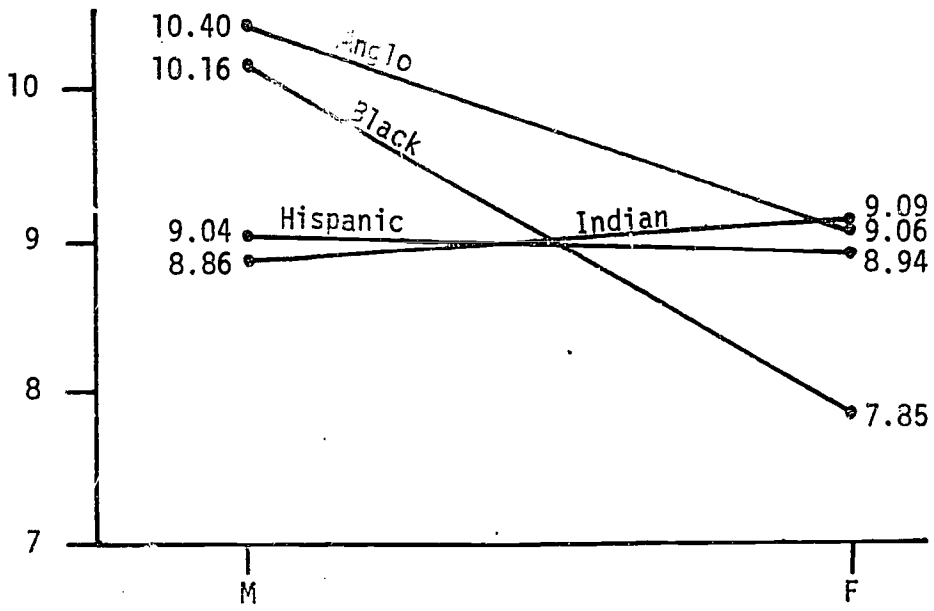

Figure 1. Testwiseness Adjusted Means of Third Grade Students

A regression analysis was carried out for the fifth grade students. There were no significant differences among ethnic groups, $F(3,258) = 1.57$. Moreover, there were no sex differences in testwiseness, $F(1,260) = .49$, and no ethnic by sex interactions, $F(3,254) = 1.19$. See Tables 6-8, and Figure 2.

Table 6. Regression of Testwiseness Scores on Ethnicity, Sex and E x S Interaction with Reading Scores as Covariate for 263 Fifth Grade Students

Source	R^2	df	F[a]	p
Covariate	.154	1,261	47.56	.000
Ethnicity	.169	3,258	1.57	.196
Ethnicity + Sex	.171	1,257	.65	.420
Sex	.158	1,260	.49	.485
Sex + Ethnicity	.171	3,257	1.62	.184
Sex + Ethnicity + E x S Interaction[b]	.183	3,254	1.19	.313

[a]Each F ratio is computed only for the last variable entered into the regression equation.

[b]Tests for covariate by ethnicity interaction, F (3,257) = 1.05 and covariate by sex interaction, F (1,259) = 2.09 were nonsignificant.

Table 7. Testwiseness Means Adjusted for Reading Scores for 263 Fifth Grade Students[a]

Ethnicity	Male	Female	Total
American Indian	9.64	8.94	9.50
Black	9.44	6.48	8.45
Hispanic	9.96	9.77	9.87
Anglo	9.54	9.51	9.52
TOTAL	9.76	9.58	9.67

[a]The regression weight (b weight) used in the adjustment process was .0566.

Table 8. Testwiseness Obtained Score Means for 263 Fifth Grade Students

Ethnicity	Male	Female	Total
American Indian	9.25	8.00	9.00
Black	8.50	5.33	7.44
Hispanic	9.77	9.52	9.65
Anglo	9.80	10.02	9.92
TOTAL	9.71	9.63	9.48

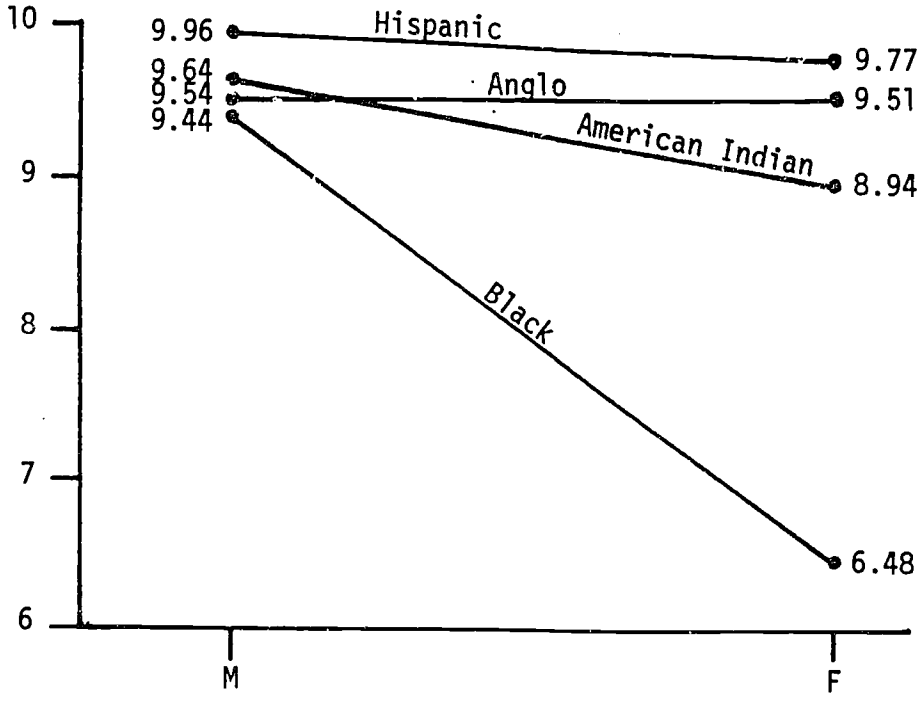

Figure 2. Testwiseness Adjusted Means of Fifth Grade Students

A regression analysis was computed for seventh grade students. Male and female students were different on testwiseness, F (1,381) = 8.49, p = .004. Ethnic differences, F (3,379) = 2.48, p = .061, and ethnicity by sex interaction were not significant, F (3,375) = 2.61, p = .051. See Tables 9-11 and Figure 3.

Table 9. Regression of Testwiseness Scores on Ethnicity, Sex and E x S Interaction with Reading Scores as Covariate for 384 Seventh Grade Students

Source	R^2	df	F[a]	p
Covariate	.062	1,382	25.18	.000
Ethnicity	.080	3,379	2.48	.061
Ethnicity + Sex	.097	1,378	7.23	.008
Sex	.082	1,381	8.49	.004
Sex + Ethnicity	.097	3,378	2.08	.102
Sex + Ethnicity + E x S Interaction[b]	.116	3,375	2.61	.051

[a]Each F ratio is computed only for the last variable entered into the regression equation.

[b]Tests for covariate by ethnicity interaction, F (3,376) = .42 and covariate by sex interaction, F (1,380) = 1.73 were nonsignificant.

Table 10. Testwiseness Means Adjusted for Reading Scores for 384 Seventh Grade Students

Ethnicity	Male	Female	Total
American Indian	8.96	10.18	9.68
Black	9.34	11.90	11.19
Hispanic	9.72	10.47	10.06
Anglo	9.92	9.99	9.95
TOTAL	9.77	10.34	10.05

[a]The regression weight (b weight) used in the adjustment process was .0372.

Table 11. Testwiseness Obtained Score Means for 384 Seventh Grade Students

Ethnicity	Male	Female	Total
American Indian	10.09	10.00	9.41
Black	9.20	11.85	11.11
Hispanic	9.49	10.32	9.88
Anglo	10.09	10.28	10.18
TOTAL	9.73	10.39	10.05

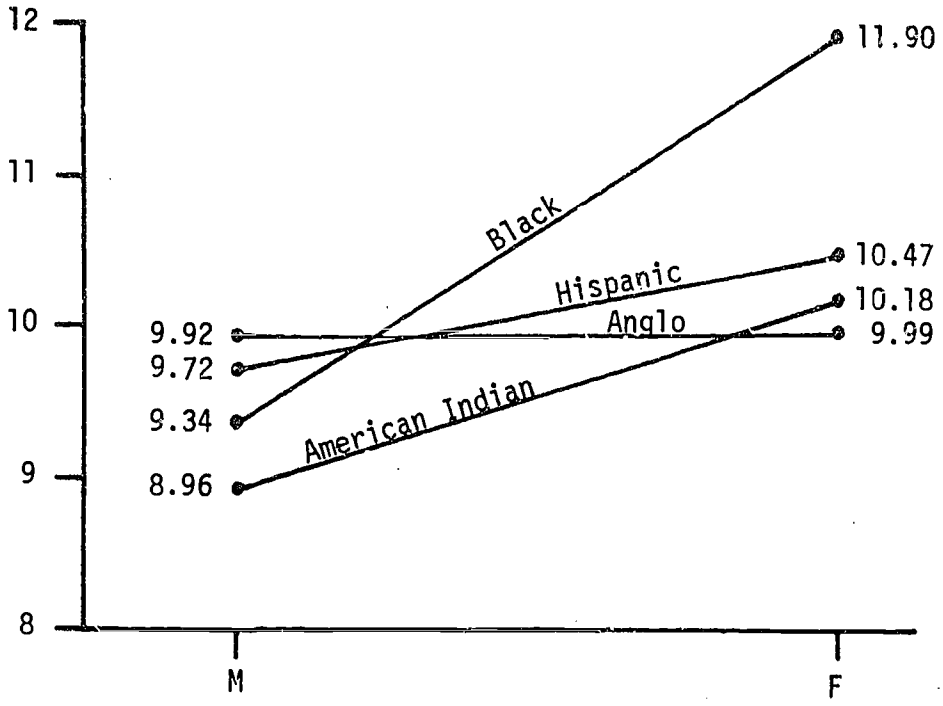

Figure 3. Testwiseness Adjusted Means of Seventh Grade Students

Results of the investigation of risktaking are contained in Appendices A-C.

A oneway ANOVA comparing these ethnic groups using testwiseness scores as the dependent variable would find significant differences at the third grade, $F(3,160) = 2.75$, $p = .045$; at the fifth grade, $F(3,259) = 3.42$, $p = .018$; and at the seventh grade, $F(3,380) = 3.01$, $p = .030$. However, when reading ability is controlled for, there were no significant differences in testwiseness among the four ethnic groups. Reading achievement accounted for a significant amount of variance in testwiseness and its control was an important factor in this research.

Using reading scores was also crucial when investigating sex differences. Without reading scores as a covariate, a simple oneway ANOVA would yield the following: third grade, $F(1,162) = 1.08$, $p = .301$; fifth grade, $F(1,261) = .081$,

$p = .777$; and seventh grade, $F(1,382) = 10.93$, $p = .001$. When reading differences were controlled, sex difference appeared at the third grade as well as the seventh grade.

The means of the groups were graphed for visual inspection. There is a difference of opinion among researchers as to the proper interpretation of significant interactions (Games, 1973; Levin and Marascuilo, 1973; Marascuilo and Levin, 1970, 1976; Betz and Gabriel, 1978). In addition, Lubin (1961) would argue that it is incorrect to report no significant ethnicity differences in the presence of a significant sex by ethnicity interaction (as we did for third grade). Rather than argue these issues at this time, we prefer to limit our paper to reporting the differences found by using reading scores as a covariate as opposed to using obtained TW scores as the dependent variable.

REFERENCES

Bajtelsmit, J. W. Development and Validation of an Adult Measure of Secondary Cue-Using Strategies on Objective Examinations: The Test of Obscure Knowledge (TOOK). Paper presented at the annual meeting of the Northeastern Educational Research Association, Ellenville, New York, October 1975a.

Betz, M. A. and Gabriel, K. R. Type IV errors and analysis of simple effects. Journal of Educational Statistics, 1978, 3 (2), 121-143.

Callenbach, C. The Effects of Instruction and Practice in Content-Independent Test-Taking Techniques Upon the Standardized Reading Test Scores of Selected Second Grade Students. Journal of Educational Measurement, 1973, 10, pp. 25-30.

Crehan, K. D., Koehler, R. A. and Slakter, M. J. Longitudinal Studies of Test-wiseness. Journal of Educational Measurement, 1974, 11, pp. 209-212.

Crehan, K. D., Gross, L. J., Koehler, R. A. and Slakter, M. J. Developmental aspects of test-wiseness. Educational Research Quarterly, 1978, 3 (1), pp. 40-44.

Cronbach, L. J. Coefficient alpha and the internal structure of tests. Psychometrika, 1951, 16, pp. 297-334.

Diamond, J. J. and Evans, W. J. An investigation of the cognitive correlates of test-wiseness. Journal of Educational Measurement, 1972, 9 (2), pp. 145-150.

Gaines, W. G. and Jongsma, E. A. The Effect of Training in Test-Taking Skills on the Achievement Scores of Fifth Grade Pupils. Paper presented at the annual meeting of the National Council on Measurement in Education, Chicago, Illinois, April 1974.

Games, P. A. Type IV errors revisited. Psychological Bulletin, 1973, 80, pp. 304-307.

Gibb, B. G. Test-wiseness as Secondary Cue Response. (Doctoral dissertation, Stanford University), Ann Arbor, Michigan: University Microfilms, 1964, No. 64-7643.

Kerlinger, F. N. and Pedhazur, E. J. Multiple Regression in Behavioral Research, New York: Holt, 1973.

Langer, G., Wark, D. and Johnson, S. Test-wiseness in Objective Tests. In PL Nacke (Ed.), Diversity in Mature Reading: Theory and Research, Vol. I, 22nd Yearbook of the National Reading Conference. National Reading Conference, Milwaukee, Wisconsin, 1973.

Levin, J. R. and Marascuilo, L. A. Type IV errors and games. Psychological Bulletin, 1973, 80, pp. 308-309.

Lubin, A. The interpretation of significant interaction. Educational and Psychological Measurement, 1961, 21, pp. 807-817.

Marascuilo, L. A. and Levin, J. R. Appropriate post hoc comparisons for interaction and nested hypotheses in analysis of variance designs: The elimination of type IV errors. American Educational Research Journal, 1970, 7, 397-421.

Marascuilo, L. A. and Levin, J. R. The simultaneous investigation of interaction and nested hypotheses in two-factor analysis of variance designs. American Educational Research Journal, 1976, 13, pp. 61-65.

Millman, J. Test-wiseness in Taking Objective Achievement and Aptitude Examinations. Final Report, 1966, College Entrance Examination Board.

Millman, J., Bishop, C. H. and Ebel, R. An Analysis of Test-wiseness. Educational and Psychological Measurement, 1965, 25, pp. 707-726.

Moore, J. C., Schutz, R. E., and Baker, R. L. The Application of a Self-Instructional Technique to Develop a Test-Taking Strategy. American Education Research Journal, 1966, 3, pp. 13-17.

Oakland, R. The Effects of Test-wiseness Materials on Standardized Test Performance of Preschool Disadvantaged Children. Journal of School Psychology, 1972, 10, pp. 355-360.

Sarnacki, R. E. An Examination of Test-wiseness in the Cognitive Test Domain. Review of Educational Research, 49(2), p. 255, 1979.

Slakter, M. J., Koehler, R. A., and Hampton, S. H. Grade Level, Sex, and Selected Aspects of Test-wiseness. Journal of Educational Measurement, 1970a, 7, pp. 119-122.

Slakter, M. J., Koehler, R. A. and Hampton, S. H. Learning Test-wiseness by Programmed Texts. Journal of Educational Measurement, 1970b, 7, pp. 247-254.

Thorndike, R. L. Reliability. In E F Lindquist (Ed.), Educational Measurement, Washington, D.C.: American Council on Education, 1951.

Woodley, K. K. Test-wiseness Program Development and Evaluation. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, Louisiana, 1973.

APPENDIX A

Reading Means Presented by Ethnicity, Grade and Sex[a]

Ethnicity	Third		Fifth		Seventh	
	M	F	M	F	M	F
American Indian	6.50	10.00	34.75	25.00	21.43	26.80
Black	15.75	11.89	25.00	21.33	28.00	30.31
Hispanic	14.50	16.49	38.30	37.10	25.72	27.73
Anglo	13.77	17.83	46.27	50.65	36.25	39.59
GRAND MEAN	15.20		41.60		31.73	

[a]Sample sizes corresponding to each mean are presented in Table 1.

APPENDIX B

Regression of Risktaking Scores on Ethnicity, Sex and E x S Interaction with Reading Scores as Covariate for 164 Third Grade Students

Source	R^2	df	F^a	P
Covariate	.005	1,162	.74	.389
Ethnicity	.027	3,159	1.24	.297
Ethnicity + Sex	.037	1,158	1.54	.217
Sex	.018	1,161	2.15	.145
Sex + Ethnicity	.037	3,158	1.04	.378
Sex + Ethnicity + E x S Interaction[b]	.085	3,155	2.70	.048

[a]Each F ratio is computed only for the last variable entered into the regression equation.

[b]Tests for covariate by ethnicity interaction, F (3,156) = .96 and covariate by sex interaction, F (1,160) = 1.25 were nonsignificant.

APPENDIX C

Regression of Risktaking Scores on Ethnicity, Sex and E x S Interaction with Reading Scores as Covariate for 263 Fifth Grade Students

Source	R^2	df	F^a	P
Covariate	.011	1,261	2.87	.091
Ethnicity	.027	3,258	1.43	.234
Ethnicity + Sex	.030	1,257	.64	.424
Sex	.012	1,260	.35	.555
Sex + Ethnicity	.030	3,257	1.53	.207
Sex + Ethnicity + E x S Interaction[b]	.074	3,254	4.07	.008

[a]Each F ratio is computed only for the last variable entered into the regression equation.

[b]Tests for covariate by ethnicity interaction, F (3,155) = 1.50 and covariate by sex interaction, F (1,159) = .82 were nonsignificant.

APPENDIX D

Regression of Risktaking Scores on Ethnicity, Sex and E x S Interaction with Reading Scores as Covariate for 384 Seventh Grade Students

Source	R^2	df	F[a]	P
Covariate	.040	1,382	15.88	.000
Ethnicity	.011	3,379	.60	.615
Ethnicity-+ Sex	.046	1,378	.48	.488
Sex	.042	1,381	.66	.416
Sex + Ethnicity	.046	3,378	.54	.655
Sex + Ethnicity + E x S Interaction[b]	.061	3,375	2.08	.102

[a]Each F ratio is computed only for the last variable entered into the regression equation.

[b]Tests for covariate by ethnicity interaction, F (3,376) = 2.08 and covariate by sex interaction, F (1,380) = 2.12 were nonsignificant.