DOCUMENT RESUME

ED 395 953 TM 025 030

AUTHOR Bennett, Randy Elliot; And Others
TITLE Sex Group Membership as a Factor in Handicapped Students' GRE General Test Performance.
INSTITUTION Educational Testing Service, Princeton, N.J.
REPORT NO ETS-RR-88-65
PUB DATE Nov 88
NOTE 21p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS College Entrance Examinations; *Disabilities; Graduate Study; Higher Education; *Performance Factors; Regression (Statistics); *Scores; *Sex Differences; *Visual Impairments
IDENTIFIERS *Graduate Record Examinations

ABSTRACT

Differences in the performance of handicapped and nonhandicapped students on the Graduate Record Examinations General Test as a function of sex group composition were investigated. Hierarchical multiple regression analyses were conducted, regressing General Test scores on sex and handicapped group membership, and adjustments to these scores were computed. Data came from 186 visually impaired students taking the regular test in a standard national administration, 151 visually impaired students taking a large type version, and 105 physically handicapped students taking the regular test type in a national administration. The resulting adjustments had only minimal effects on handicapped/nonhandicapped mean score differences, suggesting that these disparities cannot be ascribed to differences in sex group distributions. Other plausible explanations for observed score disparities between handicapped and nonhandicapped groups are discussed. (Contains 4 tables and 11 references.) (Author/SLD)

RESEARCH REPORT

SEX GROUP MEMBERSHIP AS A FACTOR IN HANDICAPPED STUDENTS' GRE GENERAL TEST PERFORMANCE

Randy Elliot Bennett
Donald A. Rock
Thomas Jirele

Educational Testing Service
Princeton, New Jersey
November 1988

Sex Group Membership as a Factor in Handicapped Students' GRE General Test Performance

Randy Elliot Bennett

Donald A. Rock

and

Thomas Jirele

Abstract

Differences in the performance of handicapped and nonhandicapped students on the GRE General Test as a function of sex group composition were investigated. Hierarchical multiple regression analyses were conducted regressing General Test scores on sex and handicapped group membership, and adjustments to these scores were computed. The resulting adjustments had only minimal effects on handicapped/nonhandicapped mean score differences, suggesting that these disparities cannot be ascribed to differences in sex group distributions. Other plausible explanations for observed score disparities between handicapped and nonhandicapped groups are discussed.

Sex Group Membership as a Factor in Handicapped Students' GRE General Test Performance

Handicapped students often perform differently from nondisabled students on admissions tests (Bennett & Ragosta, 1985). On the Scholastic Aptitude Test (SAT), for example, learning disabled examinees typically perform .5 standard deviations below the nonhandicapped group mean whereas deaf examinees perform 1 standard deviation below the mean. On the Graduate Record Examinations General Test, differences also are evident. In one analysis, physically handicapped students scored .4 standard deviations below their nondisabled peers on the Quantitative scale and .31 below on the Analytical scale (Bennett, Rock, & Jirele, 1987).

The causes of these differences in performance are not always easy to establish. Although validity studies suggest that some portion of those observed score disparities may be due to differences in developed ability (Rock, Bennett, & Kaplan, 1987; Rock, Bennett, & Jirele, 1988; Willingham, Ragosta, Bennett, Braun, Rock, & Powers, 1988), differences in the subgroups of students taking admissions tests may also play a role. For example, it is well known that the composition of some handicapped populations (e.g., those with learning disabilities) differs dramatically in sex group membership from the nondisabled population and, further, that such membership is associated with differences in test score. In national samples taking the Graduate Record Examinations

General Test during the October 1981 to June 1984 period, females typically attained scores about .6 standard deviation units lower than males on the Quantitative scale, though they achieved mean scores generally similar to males on the Verbal and on the Analytical scales (Goodison, 1983; Smith, 1984, 1985). As such, a handicapped group may appear to perform differently simply because its sex composition deviates from the base group.

Given the effect of sex differences on test performance, it is important to ask whether observed score discrepancies between handicapped and nonhandicapped groups can be accounted for by these differences. This study examined the effect of sex differences for handicapped and nonhandicapped students taking the Graduate Record Examinations General Test.

Method

Subjects

Handicapped examinees. Subjects for the study were those participating in a series of investigations of the psychometric characteristics of the General Test with handicapped students (Bennett et al., 1987; Rock et al., 1988). Samples for the series were selected from a pool of 14,142 handicapped examinees taking the General Test in domestic administrations from October 1981 through June 1984. The selection process involved four steps. First, students were selected who took General Test form C-3DGR3 in either continuously-available special administrations or nationally under standardized conditions in October 1981 or April 1982.

Form C-3DGR3 was selected because it was the only form used in both special and national administrations during this period. Approximately 1,170 disabled students took this form.

Second, to avoid confounding differences in the operation of the test owing to handicap with those due to language, 296 examinees were eliminated for whom proficiency in English was questionable. These examinees consisted of (a) noncitizens, (b) those with unknown citizenship who indicated on the GRE registration form that they communicated better in a language other than English or who failed to respond to the language question, and (c) citizens indicating that English was not their best language of communication.

Third, the remaining 873 students were classified by self-reported disability and by test administration format (e.g., visually impaired students taking the braille edition). Finally, those classifications with sample sizes equal to or greater than 100 were selected for analysis. These classifications were: visually impaired students taking the regular-type edition in a standard, national administration (n=186); visually impaired students taking large-type, extended-time administrations (n=151); and physically handicapped students taking the regular-type edition in a national administration (n=105).

Nonhandicapped examinees. A reference sample of 500 examinees was randomly drawn from the population of 20,499 individuals taking C-3DGR3 under typical testing conditions in October 1981 and April 1982. This group took the General Test

concurrently with visually impaired and physically handicapped students participating in the nationa. administration program (visually impaired students taking the large type examination received individual administrations of the form spread over the October 1981 to June 1984 period). As with the handicapped samples, students with questionable levels of English language competency were excluded from the reference groups.

Table 1 presents descriptive statistics for the four study groups. As the table shows, the reference group achieved mean scores of 498 for Verbal, 517 for Quantitative, and 514 for Analytical. With three exceptions, handicapped students performed similarly to the reference group. These exceptions were for physically handicapped students, who performed substantially lower on the Quantitative and Analytical scales, and for the visually impaired-large type group which performed considerably higher on Analytical.

With respect to sex group membership, the physically handicapped-national group had a substantial over-representation of males relative to the reference group, while the visually impaired-large type sample had a substantial underrepresentation. Given these differences, one might expect to find larger Quantitative score disparities in the physically handicapped sample and smaller ones for the visually impaired group after correcting for sex composition.

Insert Table 1 about here

Table 2 presents mean General Test scores for each group broken down by sex. As expected, the largest discrepancies between sex groups are for the Quantitative scale.

Insert Table 2 about here

Procedure

To determine the contributions of sex and handicap to total score, a hierarchical multiple regression analysis was conducted. In the analysis, the total scale score for each General Test scale was regressed on sex; then on sex and group membership based on the presence or absence of handicap; and finally on sex, group membership, and the interaction between the two. Sex group was entered into the regression before handicapped group membership because one's sex group is often determined first (i.e., at conception), and because the association between sex group membership and test score is well-documented (Goodison, 1983; Smith, 1984, 1985).

These regressions were run separately for each handicapped group combined with the reference sample. This analysis produced the proportions of variances accounted for by the variables and, in the form of the raw score regression weights, the adjusted differences in test score with selected variables held constant.

Results

Correlations between sex and handicapped/nonhandicapped group membership were nonsignificant: -.04 (two-tailed t = 1.02, df = 648, $p > .05$), for the large-type/nonhandicapped combined group, .02 (t = .52, df = 684, $p > .05$) for the visually impaired national/nonhandicapped sample, and .08 (t = 1.97, df = 103, $p > .05$) for the physically handicapped/ nonhandicapped group. These values suggest that handicap and sex are independent in these groups.

Table 3 presents the relationship between test scores and membership in sex and handicapped groups in variance terms. In the table, the predictor variables entered into the regression are listed in their orders of entry. These are followed by the proportion of variance in total score explained by each set of variables (R squared), and the increase in explained variance over that accounted for by the previous set of variables (the increment in R squared).

Insert Table 3 about here

As the table shows, even when statistically significant, the amounts of variance in total score accounted for by the independent variables are small, ranging from 0 to about seven percent. The largest increments in variance accounted for are due to sex group membership and occur for the Quantitative scale, a finding consistent with documented differences in the performance of sex groups on this scale. The increase over

sex in explained variance due to handicap is statistically significant in those cases where the scores of disabled students are most discrepant from the reference group (i.e., for physically handicapped students on Quantitative and Analytical, and for visually impaired-large type examinees on Analytical). However, the magnitude of this increase is extremely small, typically under three tenths of a percent. Finally, adding the sex-by-handicap interaction to the regression accounts for no further variance, suggesting that sex group is not more or less important in its relationship with total score for handicapped than it is for nondisabled students.

Table 4 presents unadjusted mean score differences and ones adjusted for sex group membership. Given the absence of interaction effects and the independence of the sex and handicap variables, the regression based on the main effects model (i.e., sex, handicapped group membership) was used to compute score adjustments. These adjustments were computed for the difference in mean test scores between handicapped and nonhandicapped groups controlling for sex gr up membership.

The differences between the adjusted and unadjusted scores for those handicapped groups with sex distributions that vary from the reference group are of particular interest. On Quantitative, the physically handicapped (PHN) group obtained an unadjusted score 50 points below the nondisabled mean. Because this group was disproportionately male relative to the reference sample, this score difference might be larger

given equal sex distributions. Table 4 shows, when sex was controlled, the difference between the handicapped and reference groups increased marginally to 55 points. Hence, differences in sex group membership would appear to be inflating ever so slightly the Quantitative scores of the physically handicapped sample.

Insert Table 4 about here

The second group with a sex distribution different from the reference group was the visually handicapped-large type (VIL) group. On Quantitative, this group achieved an unadjusted mean score 5 points lower than the reference sample. Since the visually impaired group was more female than the reference sample, it is possible that this minimal difference is primarily due to differences in sex distribution. As Table 4 indicates, when sex was held constant, this difference shrunk somewhat to 1 point.

Discussion

This study examined the effect of differences in sex group membership on mean score disparities between handicapped and nonhandicapped students taking the Graduate Record Examinations General Test. For the samples studied, controlling for sex resulted in only minimal adjustments, a finding consistent with attempts to adjust minority group mean scores for background differences (e.g., Holland & Thayer, 1983). These results suggest that the observed differences in

test score between handicapped and nondisabled groups were due largely to factors other than sex group membership.

What factors other than differences in sex group membership might account for the observed score differences between handicapped and nondisabled groups? Undergraduate major might seem a likely candidate because of its relationship to test score in nonhandicapped samples (Goodison, 1983; Smith, 1984, 1985), and because handicapped students taking the General Test diverge from their peers on this variable (Bennett et al., 1986). Statistical explorations of the relationship between major and test score are, however, difficult, if not impossible, to justify because of the relationship's conceptually complex nature. Although sex group membership is inherent, handicapped students (like nondisabled ones) may choose a particular undergraduate major because they are more adept at it. Just the same, they may avoid a field because their disability makes it difficult for them to succeed. As such, it is impossible to discern whether undergraduate major is a cause or effect of the observed score differences.

A second possible explanation for observed score differences is the effect of extra time (Willingham et al., 1988). This explanation is most plausible for the substantially increased performance of visually impaired-large type students on the Analytical scale. These students tend to complete this scale in greater proportions than their nonhandicapped peers (Bennett et al., 1987). As a

consequence, they have the opportunity to complete more items and obtain higher scores.

Third, the test may be measuring different attributes in handicapped and nonhandicapped groups (e.g., with physically handicapped students, speed of responding), a condition which might well result in different average score levels. Factor analytic research lends some support to this hypothesis for Analytical scores, which, for physically handicapped students and for visually impaired examinees taking the large type exam, do not appear to have a meaning comparable to nondisabled examinees (Rock et al., 1988).

Finally, some observed score disparities are likely due to real differences in the developed academic abilities measured by the test. The little research reported provides only limited support for this cause of score differences. While Verbal and Quantitative scores do seem to measure the same abilities in handicapped and nonhandicapped groups (Rock et al., 1988), preliminary research suggests that the predictive relationships of these scores to graduate grade-point average may differ from nonhandicapped samples (Braun & Ragosta, 1986). Future research might focus on replicating these predictive differences and, if replicated, determining whether they are related to the provision of excessive extra time or to other potential sources of miscalibration of the tests involved.

References

Bennett, R. E., & Ragosta, M. (1985). The technical characteristics of postsecondary admissions tests for handicapped examinees: A review of research. Journal of Special Education, 19, 255-267.

Bennett, R.E., Rock, D. A., & Jirele, T. (1987). GRE score level, test completion, and reliability for visually impaired, physically handicapped, and nonhandicapped groups. Journal of Special Education, 21, 9-21.

Braun, H. I., & Ragosta, M. (1986). The predictive validity of the GRE General Test for disabled students (RR-86-42). Princeton, NJ: Educational Testing Service.

Goodison, M. B. (1983). A summary of data collected from Graduate Record Examinations test takers during 1981-1982 (Data Summary Report #7). Princeton, NJ: Educational Testing Service, p. 172.

Holland, P. W., & Thayer, D. T. (1983). Using modern statistical methods to improve the presentation of the annual summaries of GRE candidate background data (RR-83-8). Princeton, NJ: Educational Testing Service.

Ragosta, M., & Kaplan, B. A. (1988). Views of disabled students. In W. W. Willingham, M. Ragosta, R. E. Bennett, H. I. Braun, D. A. Rock, & D. E. Powers, Testing handicapped people. Boston, MA: Allyn and Bacon.

Rock, D. A., Bennett, R. E., & Kaplan, B. A. (1987). Internal construct validity of a college admissions test across handicapped and nonhandicapped groups. Educational and Psychological Measurement, 47, 193-205.

Rock, D. A., Bennett, R.E., & Jirele, T. (1988). The Factor Structure of the Graduate Record Examinations General Test in Handicapped and Nonhandicapped Groups. Journal of Applied Psychology, 73, 383-392.

Smith, H. R. (1984). A summary of data collected from Graduate Record Examinations test takers during 1982-1983 (Data Summary Report #8). Princeton, NJ: Educational Testing Service.

Smith, H. R. (1985). A summary of data collected from Graduate Record Examinations test takers during 1983-1984 (Data Summary Report #9). Princeton, NJ: Educational Testing Service.

Willingham, W. W., Ragosta, M., Bennett, R. E., Braun, H. I. Rock, D. A., & Powers, D. E. (1988). Testing handicapped people. Boston, MA: Allyn and Bacon.

Table 1

Background Data for Study Samples

Background Characteristic	Group NH	VIL	VIN	PHN
Sample size	500	151	186	105
GRE Verbal				
Mean	498	513	486	493
SD	124	109	120	117
GRE Quantitative				
Mean	517	512	507	467
SD	135	137	131	130
GRE Analytical				
Mean	514	559	503	481
SD	131	154	135	120
Percent Male	40%	35%	43%	51%

Note. NH = nonhandicapped students taking form C-3DGR3 in October 1981 or April 1982; VIL = visually impaired students taking large-type, special administrations of C-3DGR-3 between October 1981 and June 1984; VIN = visually impaired students taking a standard administration of form C-3DGR3 in October 1981 or April 1982; PHN = physically handicapped students taking a standard administration of form C-3DGR3 in October 1981 or April 1982.

Table 2

Mean General Test Scores Broken Down

by Sex Group Membership

Group		Sex Group	
		Male	Female
NH			
	Verbal	512(120)	489(126)
	Quantitative	562(125)	487(133)
	Analytical	524(126)	508(133)
VIL			
	Verbal	508(123)	517(102
	Quantitative	546(133)	495(135)
	Analytical	565(152)	557(157)
VIN			
	Verbal	501(120)	475(121)
	Quantitative	539(142)	484(119)
	Analytical	503(136)	504(135)
PHN			
	Verbal	500(118)	490(114)
	Quantitative	486(142)	454(113)
	Analytical	480(114)	477(122)

Table 3

Relationship between GRE General Test Score and Membership in Sex and Handicapped Groups

Independent Variable	VIL/NH (n = 650) R^2	VIL/NH Increment in R^2	VIN/NH (n = 586) R^2	VIN/NH Increment in R^2	PHN/NH (n = 605) R^2	PHN/NH Increment in R^2
Verbal						
1. Sex	.004	.004	.009	.009**	.007	.007*
2. Sex & Handicap Group	.007	.003	.011	.002	.007	.000
3. Sex & Group & Sex x Group	.010	.003	.011	.000	.008	.001
Quantitative						
1. Sex	.064	.064***	.065	.065***	.055	.055***
2. Sex & Handicap Group	.064	.000	.067	.002	.078	.023***
3. Sex & Group & Sex x Group	.065	.001	.068	.001	.082	.004
Analytical						
1. Sex	.002	.002	.002	.002	.002	.002
2. Sex & Handicap Group	.022	.020***	.003	.001	.014	.012**
3. Sex & Group & Sex x Group	.022	.000	.004	.001	.014	.000

* $p < .05$
** $p < .01$
*** $p < .001$

Table 4

Mean Scale Score Differences Unadjusted and Adjusted for Sex Group Membership

Handicapped/ Nonhandicapped Difference	Combined Group		
	VIL/NH	VIN/NH	PHN/NH
		Verbal	
Unadjusted	15	-12	-5
Adjusted	16	-12	-6
		Quantitative	
Unadjusted	-5	-10	-50
Adjusted	-1	-12	-55
		Analytical	
Unadjusted	45	-11	-33
Adjusted	46	-11	-37

Note. See Table 1 note for descriptions of groups. Positive differences favor the handicapped group.