DOCUMENT RESUME

ED 395 939 TM 025 006

AUTHOR	Rivera, Charlene; Schmitt, Alicia P.
TITLE	A Comparison of Hispanic and White Non-Hispanic Students' Omit Patterns on the Scholastic Aptitude Test.
INSTITUTION	Educational Testing Service, Princeton, N.J.
REPORT NO	ETS-RR-88-44
PUB DATE	Aug 88
NOTE	24p.; Earlier version of a paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, April 1986).
PUB TYPE	Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Ability; *College Entrance Examinations; Comparative Analysis; English; High Schools; High School Students; Hispanic Americans; Item Bias; *Mexican Americans; *Puerto Ricans; Spanish; *Test Items; *White Students
IDENTIFIERS	*Hispanic American Students; Omitted Responses; *Scholastic Aptitude Test; Standardization

ABSTRACT

Standardization methodology was used to analyze omitted responses of Hispanic examinees on the Scholastic Aptitude Test. Study or focal groups were 2,956 Mexican-Americans, 3,230 Puerto Ricans, and 278,009 White test-takers. Results indicate that both Mexican-Americans and Puerto Rican students omitted fewer items than White students of comparable ability but that Mexican-Americans tended to omit even less. Examination of items with the least differential omit rate by both Hispanic groups indicate that items with true cognates or words with a common root in English and Spanish were omitted differentially less. Some dependency on ability level was observed for the omit patterns of Puerto Rican examinees. (Contains 2 tables, 6 figures, and 18 references.) (Author/SLD)

RESEARCH REPORT

A COMPARISON OF HISPANIC AND WHITE NON-HISPANIC STUDENTS' OMIT PATTERNS ON THE SCHOLASTIC APTITUDE TEST

Charlene Rivera
Alicia P. Schmitt

Educational Testing Service
Princeton, New Jersey
August 1988

A Comparison of Hispanic and White Non-Hispanic Students' Omit Patterns on the Scholastic Aptitude Test[1,2]

Charlene Rivera[3]
Development Associates

Alicia P. Schmitt
Educational Testing Service

---

[1]An earlier version of this paper was presented at the Annual Meeting of the American Educational Research Association in San Francisco, California, April 1986.

[2]Partial funding for this research was provided by Educational Testing Service and the College Board as part of the Hispanic Test Taking Skills Project.

[3]Authors are listed in alphabetical order.

Abstract

The standardization methodology was used to analyze omitted responses of Hispanic examinees on the Scholastic Aptitude Test. Results indicate that both Mexican-Americans and Puerto Rican students omitted less than White students of comparable ability but that Mexicans Americans tended to omit even less. Examination of items with the least differential omit rate by both Hispanic groups indicate that items with true cognates or words with a common root in English and Spanish were omitted differentially less. Some dependency on ability level was observed for the omit patters of Puerto Rican examinees.

Over the past twenty years educators have become increasingly concerned about the impact of tests on the educational advancement of minority students (Astin, 1982; Duran, 1983; Duran, Enright, & Rock, 1985; Linn & Harnish, 1981). For this reason there has been a concerted effort to identify the factors that affect student performance on standardized measures such as the Preliminary Scholastic Aptitude Test/National Merit Scholarship Qualifying Test, and the Scholastic Aptitude Test, called the PSAT and SAT, respectively. One factor that could affect minority student's performance is omitting or guessing. The focus for this study is to examine the differential omit patterns of Hispanic examinees utilizing existing SAT data.

A large body of literature exploring omitting or guessing strategies among minority groups does not exist. However, the handful of relevant studies that have been found used omit patterns as an indirect measure of guessing. Although omitting may be a reflection of guessing, it cannot be interpreted solely from this perspective. When items are answered or not omitted, three options are possible--an individual knows the answer and responds with certainty; an individual does not know an answer with certainty, but can venture an educated guess; and an individual has no knowledge base out of which to operate and so randomly selects a response or guesses.

Pike (1979) defines guessing as "answering a test question in the absence of certainty as to the correct response" (p. 7). He adds that guessing patterns may be divided into three categories: "guessing that is blind or random, guessing that is spurious or based on a hunch, and guessing that is based on partial information" (1979, p. 7).

Whether or not an individual guesses tends to be an individual characteristic. As Pennock-Roman (1984) indicates, the willingness to take a risk and guess may decrease with experience and age. Among minority groups, the

research that has been conducted suggests that differences in guessing patterns exist, particularly for different item formats. For example, Flaugher and Pike (1970), in a study using the PSAT, found that inner city Blacks tended to answer in a more random fashion than the norming population in all areas of the verbal subtest except analogies. The results of this study suggest that students may use different guessing strategies, depending upon item type.

In another study of analogy items, Pike (1980) used data from the Graduate Record Examination (GRE) and found that

> . . . intergroup differences [to tests and to test-like tasks] do not exist . . . There is no real evidence that the minority groups here studied were different from their majority counterparts in any way not explained by overall test performance itself. Minorities score less well. But minorities do not show test performances . . . which are different from low scoring majority groups (pp. 78-79).

The only possible difference found by Pike was that the average performance of Chicana females who omitted an item was lower than that of matched White comparison females who omitted the same item. Pennock-Roman (1984) critiques the study. Specifically, she cites the ambiguity of the index used by Pike to arrive at the above stated conclusion. She states that the results are paradoxical. On the one hand, low ability Chicanas may guess less because they are less adventuresome, while on the other hand, lower scores for Chicana females may be due to a higher overall omission rate.

Since it was not possible to conduct meta-cognitive research where test-takers could provide a self-report of why they responded to or omitted an item, for this exploratory study, we compared the omit strategies of Hispanic and White students. These were assessed by comparing the proportion of items omitted at each scaled score level on a recently disclosed form of the verbal section of the

SAT. Using this approach, we examined omitted items on the SAT to infer why students responded differentially to certain items.

METHOD

Mexican-American and Puerto Rican students who took the same verbal SAT test form and who indicated that English was their best language on the Student Descriptive Questionnaire (SDQ) were compared with White students who also indicated that English was their best language. For the study, this background variable was used as a sample selection criterion to control for overall language differences among the students. Specifically, the study or focal groups were 2,956 Mexican-Americans, 3,230 Puerto Ricans and 278,009 White test-takers. It should be noted that only Puerto Ricans residing in the continental United States were included in the sample.

The SAT verbal test taken by the study students consisted of 85 verbal items. Instructions given to students prior to taking the test indicate that it is appropriate to guess if an educated choice can be made by eliminating responses that are definitely wrong. Students are informed that they will be credited for each question answered correctly but that their total score will be adjusted for each incorrect answer (Taking the SAT, 1983).

For this study, the statistical standardization method devised by Dorans and Kulick (1983; 1986) was revised to analyze omitted rather than correct student responses on the SAT. The method was originally devised to identify unexpected differences in item performance after controlling for ability level. Because of this the standardization procedure was deemed an appropriate method to compare the omit differences betwe n Hispanic and White respondents of equal ability. Through this method conditional estimates of observed percent omit were computed

across the entire scaled score range for each item for both the base and the focal or study groups.

The responses used in the standardization procedure were redefined to be omitted responses rather than correct responses. Differences in probabilities given by this index are referred to as "unexpected" because they identify discrepancies between groups with the same score on a specific ability measure and among whom one would not expect differences.

This index provides a numerical value of the difference in the conditional probability of omitting the item across total scaled score levels. Mathematically, this difference (D_{STD}) is defined as:

$$D_{STD} = \sum_{s=1}^{S} N_{fs} \, (O_{fs} - O_{bs}) \; / \; \sum_{s=1}^{S} N_{fs}$$

where S is the number of score levels, N_{fs} is the number of individuals at score level s in subpopulation f, (focal group) and O_{fs} is the conditional probability of omission in subpopulation f at score level s. O_{bs} is the proportion of examinees in the base group who omitted the item at score level s. If there is no difference in conditional probabilities of omitting an item, D_{STD} should equal zero. When a focal or study group omits an item at a greater rate than expected, D_{STD} will be positive. When a focal group omits an item less than expected, D_{STD} will be negative. A value of $|D_{STD}|$ greater than or equal to .05 has been used as a criterion to identify differences that merit further investigation during research (Rogers, Dorans, & Schmitt, 1986; Dorans & Kulick, 1986).

For this study, the relationship between the percent omits and the SAT scaled scores were examined for those items where unexpected omit patterns occurred. The patterns found for Whites were compared independently with each Hispanic subgroup. Since White SAT takers comprised the largest group, they

provided the most stable estimate of the conditional probability of omitting across the entire scaled score range. For this reason, they were selected to be the base or comparison group for the study. The Mexican-American and Puerto Rican students were defined as the focal groups for whom item omit probabilities were estimated. Items were considered to exhibit differential item omission rates when the probability of omitting or not omitting an item was greater for the focal group than for the base group after ability was taken into consideration.

RESULTS AND DISCUSSION

The analyses provided independent comparative data for Mexican-American and White students and for Puerto Rican and White students on each of the 85 verbal SAT test items. The results indicated that both groups of Hispanic students omitted less than expected when compared to White students. Table 1 presents the comparison of omitted items for White and Hispanic students by item type. Item types included in the SAT-Verbal test are antonym, sentence completion, reading comprehension, and analogy. Descriptive statistics are given for the standardized difference index across total item type. Items were classified as positively or negatively differential. Positive items were described as those where the study group as compared to the base group omit more than expected while negative items were inversely defined. The numbers in parentheses indicate discrepant items. They were flagged using a criterion of $|D_{STD}|$ greater than or equal to .05.

Insert Table 1 about here

Mexican-American students seemed to omit less than expected for all item types except analogies as observed by the proportion of items with negative to positive D_{STD} omits. They omitted differentially less than expected according to the D_{STD} criterion on 4 antonym and 1 analogy item. No items were flagged as being differentially omitted more. Puerto Rican students had more items where they omitted less than expected on antonyms and reading comprehension item types and more items where they omitted more than expected on sentence completion and analogies. They omitted differentially less than expected according to the D_{STD} criterion of $|D_{STD}|$ greater than or equal to .05 on one antonym item and more than expected on one analogy item.

A graphical display of the results presented on Table 1 are found in Figures 1 and 2. The standardized difference of percent omits or the D_{STD} is represented on the abscissa of these plots. The standard deviation of the standardized difference is shown on the ordinate axis. Each item type is identified by different symbols. Since the criterion used to flag items with extreme standardized differences is a $|D_{STD}|$ greater than or equal to .05, dotted perpendicular lines were drawn at intervals of .05. Figure 1 presents the difference in percent omit results for the Mexican-American group and Figure 2 presents the findings for Puerto Rican students.

Insert Figures 1 and 2 about here

Results in these figures indicate that Mexican-Americans omitted less than expected on all item types except analogies. Puerto Ricans omitted less than expected on antonym and reading comprehension items while they omitted more than expected on sentence completion and analogy items.

Items flagged as having positive or negative unexpected differential omit rates for either Hispanic group are presented in Table 2. The differential item functioning for the correct response[3] and the omit D_{STD} are listed for Mexican-American and Puerto Rican students independently. The last column indicates, for each item, the group that had the lowest conditional omit rate. These items were examined to assess commonalities that might explain differences.

While five items were flagged because they had unexpectedly extreme lower omit rates for the Mexican-American students, only one of the same five items was flagged for the Puerto Rican students. Content analysis indicated that each of the 5 items contained true cognates or words in the item that have a common root in Spanish and English. Cognates were located in at least the stem, key, and/or distractors. The one antonym item (FACILITATE:complicate) flagged for both groups had cognates in every item component. Another item with all components as cognates (APROCRYPHAL:AUTHENTICATION) was flagged for the Mexican-American group only.

Insert Table 2 about here

Inspection of the standardized differential item functioning value for the correct response on these items revealed that the antonym item flagged for both Hispanic groups was differentially easier for Puerto Ricans and Mexican-Americans than it was for the White test-takers of comparable ability. This finding, also supported by other research (Alderman, 1981; Breland, Stocking, Pinchak, & Abrams, 1976; Schmitt, 1985; 1987; 1988), led us to hypothesize that for items containing true cognates in all components of the item, Hispanic students may be using their linguistic and conceptual knowledge of the Spanish language to respond and in consequence omitting differentially less.

---

[3] For another study, Schmitt (1987) computed the standardized differential item functioning for the items analyzed here.

The last analogy listed--GULLY:EROSION::MINE:EXCAVATION--was the only item flagged that had a significantly larger omit rate for the Puerto Rican group. Although the Mexican-American omit rate for this item was not larger than an D_{STD} of .05, it was also the item that was omitted the most by this group. One possible explanation for the unexpected omission of this item by both Hispanic groups and especially among Puerto Ricans might be regional differen in the frequency of use of the word "gully". Notice, in addition, that this item was also differentially harder for the Puerto Rican group.

The six flagged items were analyzed further by plotting the D_{STD} of percent omits conditional on percent correct for both the Hispanic groups. Figures 3 and 4 provide examples of the resulting plots for the first analogy item for the Mexican-American and Puerto Rican groups, respectively. This item had true cognates in the stem, key, and distractors. As can be observed in the plots, the points indicating D_{STD} in percent omits for each Hispanic group tend to fall below the horizontal line indicating a zero difference and thus reflect that for both groups the item was omitted less than it was by comparable White examinees.

Insert Figures 3 and 4 about here

An interesting finding was observed when items where Mexican-American and Puerto Rican students' omit patterns differed. Mexican-American students, at all ability levels, were found to omit less than White students. They also omitted less than Puerto Ricans. For example, at all ability levels, Mexican-American students responded more than comparable White students to the word "table", meaning to delay action, and to the antonym item with the word "acerbic". For both items, the item difficulty for the Mexican-Americans, .26 and .14, and for the White students of comparable ability, .32 and .17, respectively, indicates

that the items were quite difficult for either group. Although the difficulty level for these items, .25 and .13, respectively, for the Puerto Rican group, was similar to that of Mexican-American students, the Puerto Rican students' omit patterns varied more according to ability level.

Figures 5 and 6 for Mexican-American and Puerto Rican students, respectively, display examples of how differential omit patterns might be dependent on total scaled score. As illustrated in Figure 6, Puerto Rican students at the middle-ability level (scaled scores ranging from approximately 350 to 650), tended to omit more than White and Mexican-American students. At the lower ability level (scaled scores lower than approximately 350), they, like Mexican-American students, tended to omit less than White students. The omit pattern for both the Mexican-American and the Puerto Rican students at the higher ability range (scaled scores higher than approximately 650), tended to be very similar to that of White test-takers. This observation suggests that Puerto Ricans at the middle-ability level may realize when they do not know the answer to an item and are thus more discerning, or that they may be more inhibited than Mexican-American students in responding to moderately-difficult and difficult items.

Insert Figures 5 and 6 about here

Since the results presented are based on a limited number of items they should be considered tentative. In addition, inferences made regarding the dependency of omit rates on total scaled scores should be viewed cautiously, particularly because of the relatively small sample sizes at the higher scaled score ranges.

CONCLUSIONS

Both Hispanic study groups who took the SAT verbal test tended to omit less than White student of comparable ability. This finding was more salient for the Mexican-American study group suggesting that, overall, students in this group may be omitting less than Puerto Rican and White test-takers of comparable ability.

Pike's (1980) findings that Chicana females who omit more are a less able subgroup of the population was not substantiated in this study. Although analyses were not conducted by sex, overall, the Mexican-American students' omit patterns were not dependent on ability differences. An interesting finding of this investigation was that for some items where the omit response of Mexican-Americans differed from Puerto Rican omit patterns, the omit patterns for the Puerto Rican group did seem to be dependent upon ability. The Puerto Rican group that omitted more scored at the middle ability range; fewer omits were found for students who scored at the lower ability levels.

Content analyses of items showing negative differential omit rates indicated that true cognates were present in these items. When an item had several true cognates, it was omitted less by both groups of Hispanic students. The use of true cognates has been found to positively affect Hispanic student's performance (Alderman & Holland, 1981; Breland, et al., 1976; Schmitt, 1985; 1987; 1988). The discrepancy in omit patterns for both Mexican-American and Puerto Rican study groups tended to be in the same direction. Examination of the only item flagged as discrepant for both study groups showed that the presence of true cognates might have influenced the lower omit pattern found.

Because this study was exploratory in nature and omit discrepancies were found in only a limited number of items, our interpretation of the results should be followed up by further study. One approach would be to replicate this study

by examining other verbal standardized tests or other forms of the SAT. Another approach would be to create an experimental testing situation where cognate content was manipulated and metacognitive explanations collected.

Although tentative, these findings are of interest to test developers, researchers, and practitioners. Results indicating that Hispanics have lower omit rates than White students suggest that both groups may have different approaches to taking tests. In addition, the use of cognates by Hispanic students indicates that language background can influence how they approach certain items. Moreover, these results also indirectly provide generally useful information about omitting strategies used by Hispanic students. From a practical perspective, these findings have direct implications for training Hispanic students in test-taking strategies. More generally, they generate hypotheses worthy of further exploration by psychometricians and researchers.

/kad
AS\RRACOMP

REFERENCES

Alderman, D. L., & Holland, P. W. (1981). Item performance across native language groups on the test of English as a Foreign Language. (TOEFL RR 9). Princeton, NJ: Educational Testing Service.

Astin, A. W. (1982). Minorities in higher education. London, England: Jossey-Bass Publishers.

Breland, H. M., Stocking, M., Pinchak, B. M., & Abrams, N. (1976). The cross-cultural stability of mental test items: An investigation of response patterns for ten sociocultural groups. (PR-74-2). Princeton, NJ: Educational Testing Service.

Dorans, N. J., & Kulick, E. (1983). Assessing unexpected differential item performance of female candidates on SAT and TSWE forms administered in December 1977: An application of the standardization approach. (RR-83-9). Princeton, NJ: Educational Testing Service.

Dorans, N. J., & Kulick, E. (1986). Demonstrating the utility of the standardization approach to assessing unexpected differential item performance on the Scholastic Aptitude Test. Journal of Educational Measurement, 23, 355-368.

Duran, R. P. (1983). Hispanics' education and background: Predictors of College Achievement. New York: College Entrance Examination Board.

Duran, R. P., Enright, M. K., & Rock, D. (1985). Language factors and Hispanic Freshmen's student profile. New York: College Entrance Examination Board.

Flaugher, R. L., & Pike, L. W. (1984). Reactions to a very difficult test by an inner city high school population: A test and item analysis. RM-70-11. Princeton, NJ: Educational Testing Service.

Linn, R. L., & Harnisch, D. L. (1981). Interactions between item content and group membership on achievement test items. Journal of Educational Measurement, 18, 109-118.

Pennock-Roman, M. (1984. April). A reanalysis of implicit guessing strategies on the GRE. Paper presented at the Annual Meeting of the American Educational Research Association, Montreal, Canada.

Pike, L. W. (1980). Implicit guessing strategies of GRE-aptitude examinees classified by ethnic group and sex. (GRE No. 75-10P). Princeton, NJ: Educational Testing Service.

Pike, L. W. (1979). Short-term instruction, testwiseness, and the scholastic aptitude test: A literature review with research recommendations. New York: College Entrance Examination Board.

Rogers, J., Dorans, N. J., & Schmitt, A. P. (1986). Assessing unexpected differential item performance of Black candidates on SAT form 3GSA08 and TSWE form E43. (SR-86-22). Princeton, NJ: Educational Testing Service.

Schmitt, A. P. (1985). Assessing unexpected differential item performance of Hispanic candidates on SAT form 3FSA08 and TSWE form E47. (SR-85-169). Princeton, NJ: Educational Testing Service.

Schmitt, A. P. (1987). Unexpected differential item performance of Hispanic examinees. In A. P. Schmitt & N. J. Dorans (Eds.) Differential item functioning on the Scholastic Aptitude Test (RM-87-1). Princeton, NJ: Educational Testing Service.

Schmitt, A. P. (1988). Language and cultural characteristics that explain differential item functioning for Hispanic examinees on the Scholastic Aptitude Test. Journal of Educational Measurement, 25, 1-13.

Taking the SAT: A guide to the Scholastic Aptitude Test. (1983). New York: College Entrance Examination Board.

Wright, D. (1987). Empirical comparison of the Mantel-Haenszel and standardization methods of detecting differential item performance. In A. P. Schmitt & N. J. Dorans (Eds.) Differential item functioning on the Scholastic Aptitude Test (RM-87-1). Princeton, NJ: Educational Testing Service.

Table 1

Standardized Difference on SAT-V Omitted Items
for Hispanic and White Examinees

	Mexican-American/White			Puerto Rican/White		
	# Items[a]	D_{STD}		# Items[a]	D_{STD}	
		Mean	SD		Mean	SD
Antonym						
Total	25	-.02	.02	25	-.00	.02
Positive	9 (0)	.01	.01	9 (0)	.01	.01
Negative	16 (4)	-.03	.02	16 (1)	-.01	.02
Sentence Completion						
Total	15	-.01	.02	15	.00	.00
Positive	3 (0)	.00	.00	14 (0)	.00	.00
Negative	12 (0)	-.01	.02	1 (0)	-.00	---
Reading Comprehension						
Total	24[b]	-.01	.01	24[b]	-.00	.01
Positive	2 (0)	.00	.00	9 (0)	.00	.00
Negative	22 (0)	-.01	.01	15 (0)	-.00	.00
Analogy						
Total	19[b]	-.01	.03	19[b]	.01	.03
Positive	10 (0)	.01	.01	13 (1)	.02	.03
Negative	9 (1)	-.03	.02	6 (0)	-.02	.02
Total Test						
Total	83[b]	-.01	.02	83[b]	.00	.02
Positive	24 (0)	.01	.01	45 (1)	.01	.02
Negative	59 (5)	-.02	.02	38 (1)	-.01	.01

Note. Standardized Difference (D_{STD}) is an index of the discrepancy of the conditional probability of omitting the item across total score levels.

[a]Values in parentheses identify number of positive or negative discrepant items using a criterion of $|D_{STD}| \geq .05$.

[b]One Reading Comprehension item and one Analogy item were excluded from the calculation because their percent correct was less than .05.

Table 2

SAT-Verbal Items with High Standardized Differential
Omit and Difficulty Rates for Hispanic Examinees

Item[a]	Mexican-American/White D_{STD}		Puerto Rican/White D_{STD}		Omitted Less than Expected
	DIF[b]	Omits	DIF[b]	Omits	
Antonym Items					
FACILITATE: (A) intensify (B) mobilize (C) decline (D) complicate (E) meditate	.13	-.06	.19	-.06	Mexican-Americans & Puerto Ricans
CAPRICIOUS: (A) feeble (B) scant (C) constant (D) heavy (E) ungainly	.01	-.05	.01	-.03	Mexican-Americans
ENDEMIC: (A) alien (B) bloated (C) energetic (D) virulent (E) angelic	-.00	-.05	.01	-.01	Mexican-Americans
CLEMENCY: (A) foolishness (B) sadness (C) loudness (D) restlessness (E) harshness	.04	-.05	.07	-.03	Mexican-Americans
Analogy Items					
APOCRYPHAL:AUTHENTICATION:: (A) ridiculous:familiarity (B) detrimental:intention (C) mystical:intolerance (D) dogmatic:impatience (E) profane:sanctity	.03	-.07	.04	-.03	Mexican-Americans
GULLY:EROSION:: (A) drought:precipitation (B) mine:excavation (C) clot:dispersion (D) forest:cultivation (E) water:inundation	-.03	.03	-.09	.10	Whites

[a] True cognates are underlined.

[b] Differential item functioning for the correct answer was obtained from analyses conducted by Schmitt (1987).

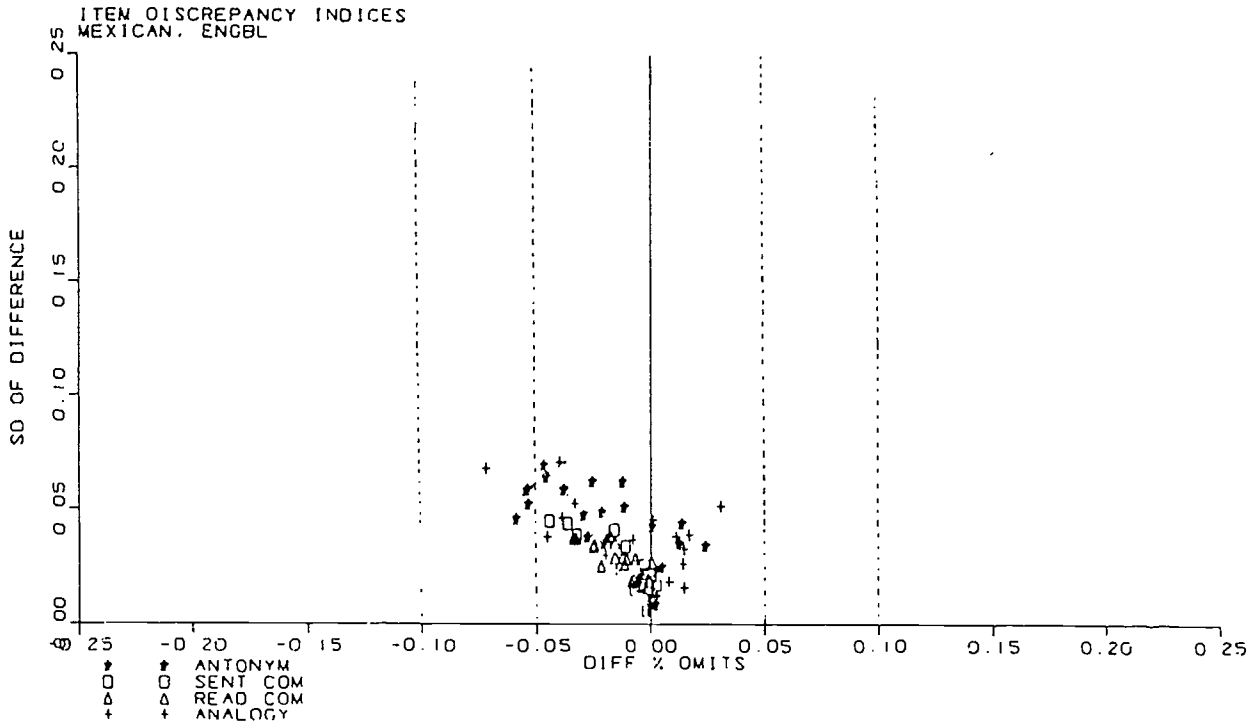

Figure 1. Standardized Difference of Percent Omits on the SAT-V for Mexican-American Examinees.

Note. Perpendicular dotted lines are drawn at |.05| intervals for the standardized difference.

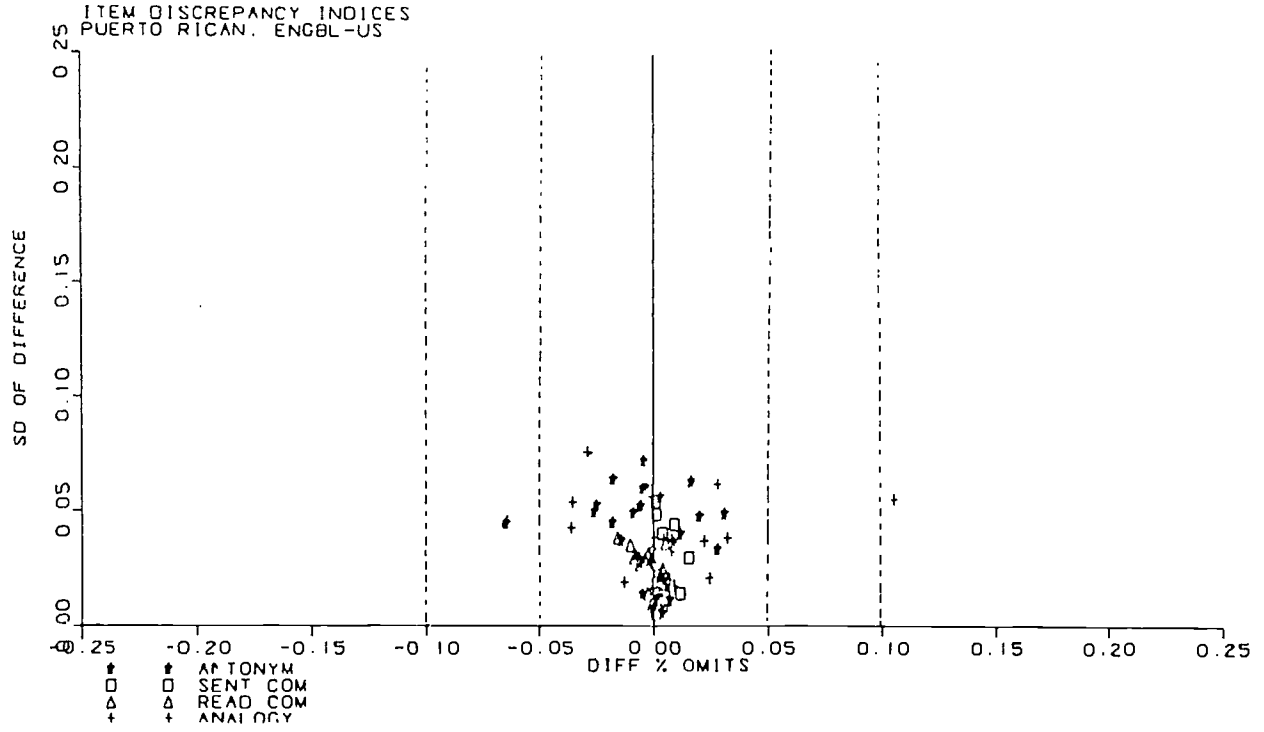

Figure 2. Standardized Differences of Percent Omits on the SAT-V for Puerto Rican Examinees.

Note. Perpendicular dotted lines are drawn at |.05| intervals for the standardized difference.

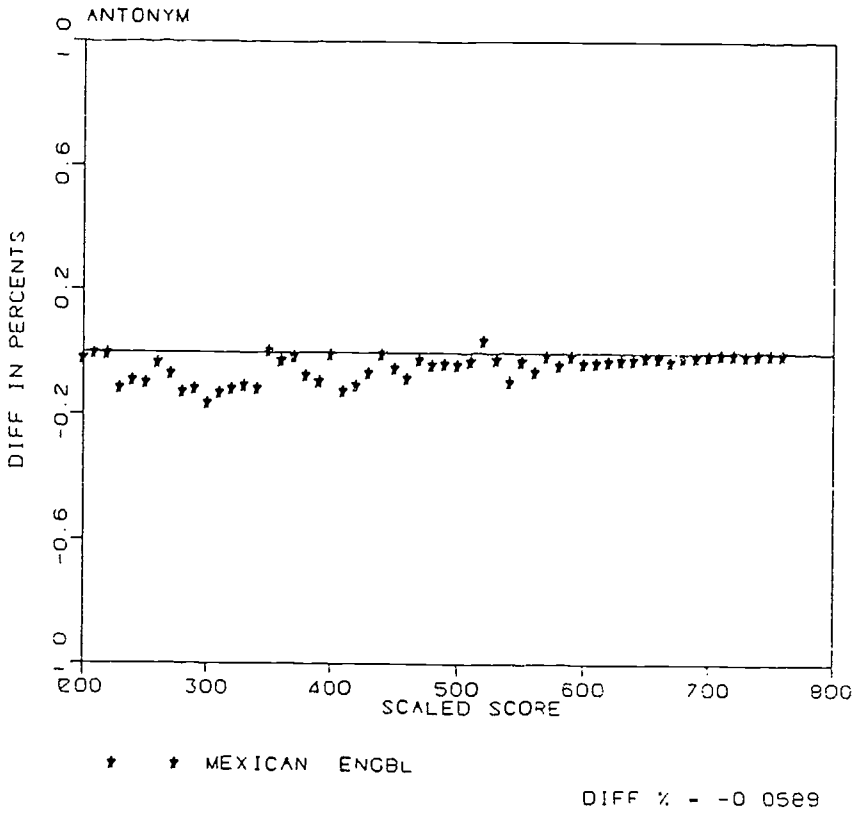

Figure 3. Conditional Probability Plot of Omit Rate for Mexican-Americans on an item with cognates.

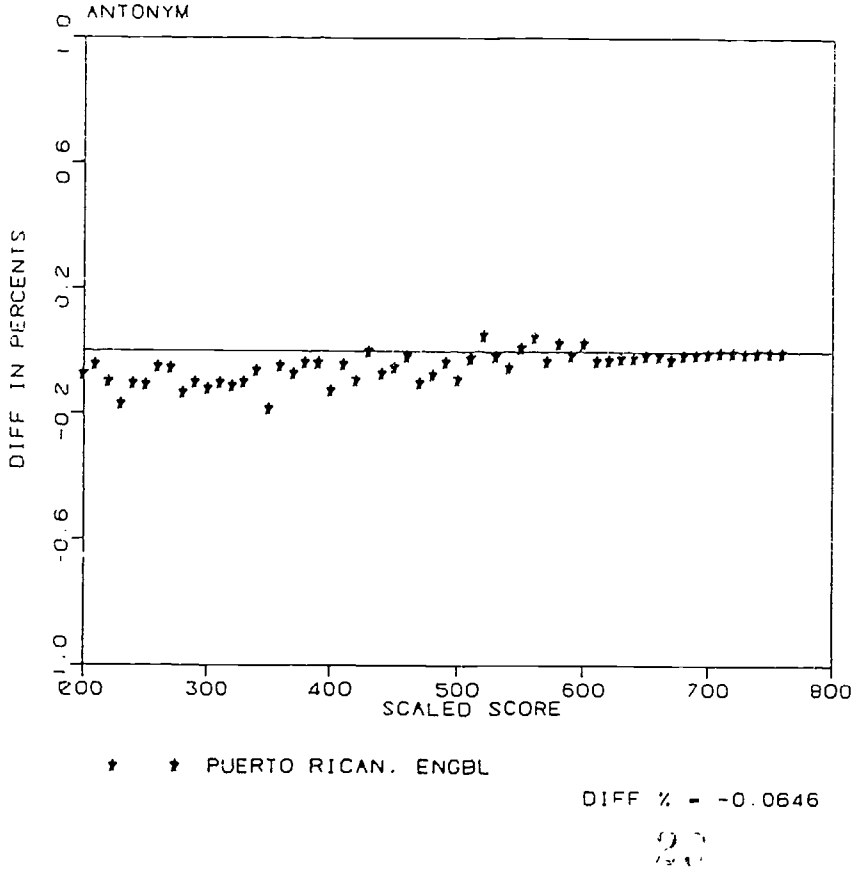

Figure 4. Conditional Probability Plot of Omit Rate for Puerto Ricans on an item with cognates.

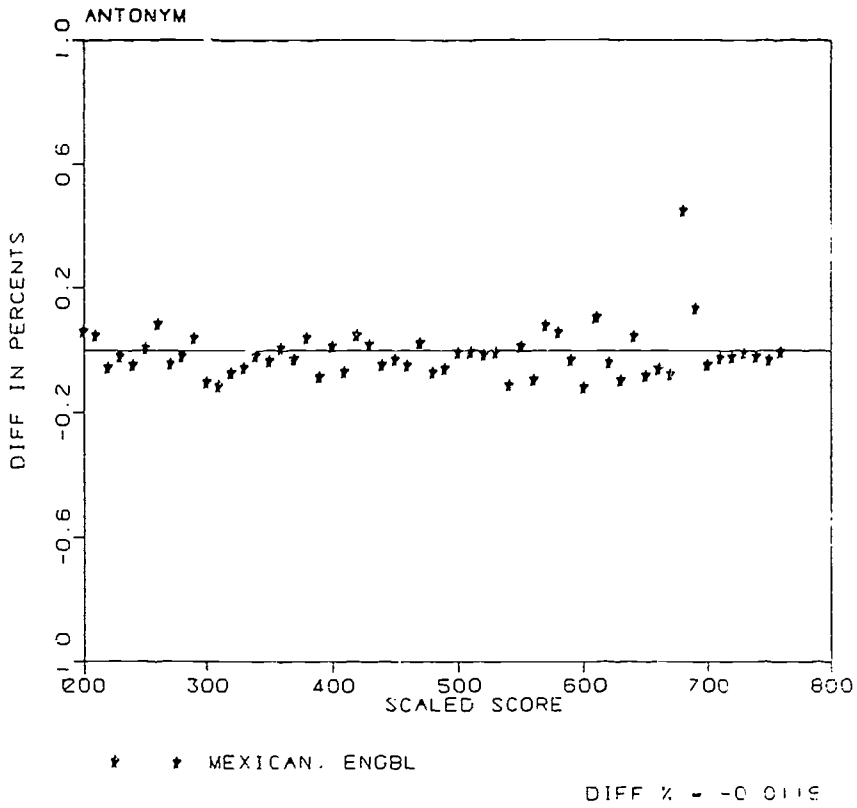

Figure 5. Conditional Probability Plot of Omit Rates for Mexican-Americans

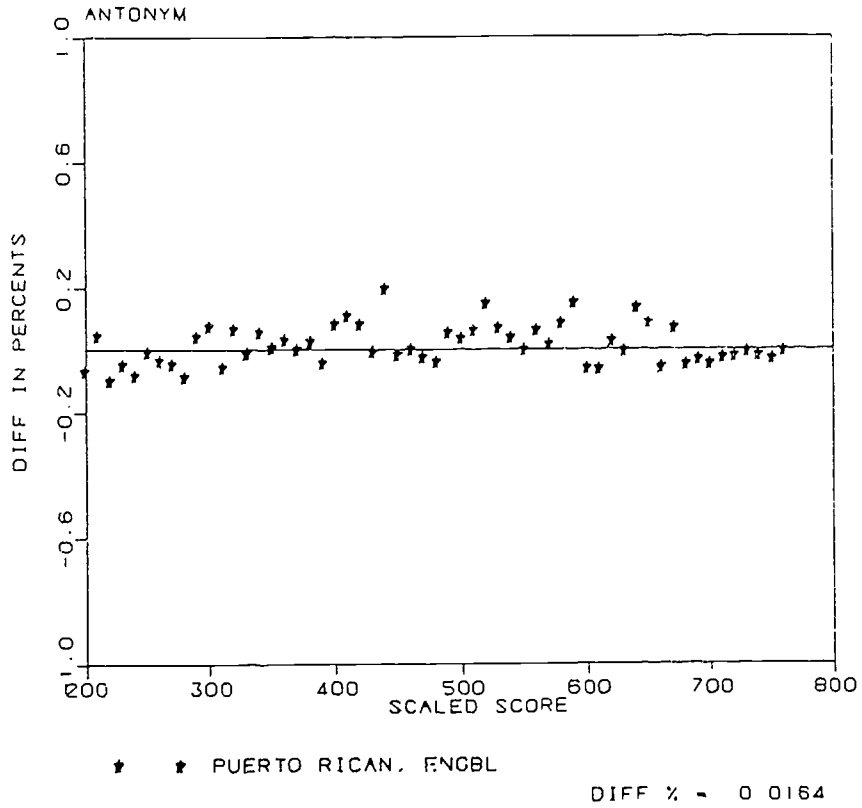

Figure 6. Conditional Probability Plot of Omit Rates for Puerto Ricans